DTU Library

Integrated control platform for converged optical and wireless networks

Yan, Ying

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

Integrated Control Platform for Converged Optical and Wireless Networks

Ying Yan

December 2009

Networks Technology & Service Platforms
Department of Photonics Engineering
Technical University of Denmark
2800 Kgs. Lyngby
DENMARK

To my dear parents: Zhihua Yan and Zhiying Niu.

Abstract

The next generation of broadband access networks is expected to be heterogeneous. Multiple wired and wireless systems can be integrated, in order to simultaneously provide seamless access with an appropriate Quality of Service (QoS). Wireless networks support ubiquitous connectivity yet low data rates, whereas optical networks can offer much higher data rates but only provide fixed connection structures. Their complementary characteristics make the integration of the two networks a promising trend for next generation access networks. With combined strengths, the converged network will provide both high data rate services and connectivity at anytime and anywhere.

One major challenge in the interworking is how to achieve seamless integration. There are many aspects involved in designing an integrated control platform, such as QoS provisioning, mobility, and resiliency. This dissertation introduces the complementary characteristics of the optical networks and the wireless networks, addresses motivations for their interworking, discusses the current progress in hybrid network architectures as well as the functionalities of a control system, and identifies the achieved research contributions in the integrated control platform design.

To achieve an integrated and unified control platform, enhanced signalling protocol plays an important role in gluing the two different technologies. Consequently, an integrated resource management system is developed. Furthermore, an admission control scheme for connections in the wireless domain can be jointly designed with the optical upstream bandwidth allocation scheme in the optical domain. Higher resource utilization is achieved due to an effective manipulation of the overall resources of two networks.

In the converged optical and wireless network scenario, multiple wireless networks are adjacent to the backbone optical network. Although the local resource allocation mechanisms implemented in the wireless networks individually can provide certain levels of QoS provisioning, proper load balancing and resource allocation schemes are needed in order to utilize the integrated resources effectively and efficiently. An integrated load balancing mechanism is proposed to take advantage of the centralized control in the optical network. A modified signalling protocol is developed to improve information exchanged between optical and wireless domains. Traffic load and network resources are distributed based on the network states, channel conditions, and QoS requirements.

A new aspect in the design of future network is the energy efficiency. An energy management mechanism is proposed and evaluated for the optical network. With regard to power saving, a sleep mode operation is developed. Therefore, power is conserved by switch off some operating functions. The sleep period and wakeup period are computed and assigned using two alternative scheduling schemes, which show trade-off performances on energy efficiency, queuing delay and network bandwidth utilization.

To summarize, this dissertation presents new knowledge by developing a novel integrated control platform for the converged optical and wireless network. Several contributions are presented by investigating network architectures, protocols, and energy issues to obtain efficient hybrid networks.

Resumé

Den næste generation af bredbåndsnetværk forventes at være heterogene. Flere kablede og trådløse systemer kan integreres og samarbejde om at levere en forbindelse med en passende Quality of Service (QoS). Optiske netværk kan tilbyde meget høje datahastigheder, men kræver en fast tilslutning, hvorimod trådløse netværk er mindre afhængige af nodernes geografiske placering, men opererer ved meget lavere hastigheder. Deres komplementære egenskaber gør integrationen af de to netværkstyper oplagte for næste generations tilgangsnetværk. Det resulterende netværk kombinerer styrkerne ved de to og tilbyder høj datahastighed samt mulighed for tilslutning når som helst og hvor som helst.

En stor udfordring i det indbyrdes samarbejde er, hvordan man opnår problemfri integration. Der er mange aspekter involveret i udarbejdelsen af en integreret kontrol platform, såsom QoS provisionering, mobilitet, og robusthed. Denne afhandling gennemgår de forskellige karakteristikker ved optiske og trådløse netværk. Derudover diskuteres motivationen for, og den nuværende status for udviklingen af, sådanne hybride netværksarkitekturer, herunder kontrolfunktionaliteten, og de opnåede forskningsresultaters bidrag til designet af den integrerede kontrolplatform identificeres.

For at opnå en integreret og ensartet kontrolplatform, er det nødvendigt med udvidede signaleringsprotokoller ofr at sammenkoble de to forskellige teknologier. Derfor er et integreret ressourcestyringssystem blevet udviklet.

Herudover er udviklet en kontrolmetoder, der sammenkæder en kontrolmetode for forbindelser i det trådløse domæne med uplink båndbreddeallokerings-metoden i optiske netværk. Herved opnås bedre ressourceudnyttelse, ved effektiv styring af de samlede ressourcer i de to netværk.

I det sammensmeltede optiske og trådløse netværksscenarie, arbejder flere trådløse netværk sideløbende med det optiske kernenetværk. Selvom de lokale ressourcetildelingsmekanismer, der fungerer på det trådløse netværk individuelt kan tilbyde visse niveauer af QoS provisioning, er overordnet ressourcetildeling og udjævning af belastningsgraden på de enkelte forbindelser en nødvendighed for at udnytte de samlede ressourcer effektivt. Til balancering af trafikbelastningen foreslås en mekanisme, som drager fordel af den centrale kontrol i det optiske netværk, og en modificeret signaleringsprotokol er udviklet til at forbedre udvekslingen af oplysninger mellem optiske og trådløse domæner. Trafikbelastning og netværksressourcer fordeles baseret på netværkstilstande, kanal betingelser og QoS krav.

Et nyt perspektiv i udformningen af fremtidige netværk er energieffektivitet. I denne sammenhæng er en energistyringsmekanisme til optiske netværk blevet udviklet og evalueret. Dette inkluderer blandt andet en dvaletilstandsoperation, som sparer strøm ved at slukke visse systemfunktioner når de ikke benyttes. Dvaleperioden og oppeperioden er beregnet og tildelt ved hjælp af to forskellige planlægningsmekanismer, hvilket viser de forskellige kompromiser mellem energieffektivitet, kø forsinkelse og udnyttelsen af netværkets båndbredde.

For at opsummere, præsenterer denne afhandling ny viden ved at udvikle en ny integreret kontrol platform for integrerede optiske og trådløse netværk. Herudover undersøges netværksarkitekturer, protokoller og energioptimeringer for at opnå de mest effektive hybridnetværk.

Acknowledgements

The process of achieving the Ph.D. is a wonderful journey in my life.

My mentor, Dr. Lars Dittmann, has been my greatest inspiration. His creativity, dedication, and confidence have guided my thinking and attitude towards work and life. He has always taught me to think positively. Over these years, he has steered me through the challenge with his wisdom and his encouragement. I thank him for giving me the opportunity and keeping faith with me.

I deeply admire Dr. Leoneid Kazovsky from Photonic Network Research Laboratory (PNRL) at Stanford University, for his help and encouragement. Working as a visiting scholar in his research group is a pleasant experience. He has encouraged me to think novelly, and his philosophy toward work has taught me how to innovate.

I appreciate the help and support from my Ph.D. committee members - Dr. Michael S. Berger and Dr. Henrik Wessing. I have greatly benefitted from their insights through research inputs. Special thank goes to our group members: Dr. Villy B. Iversen, Dr. Lars Staalhagen, Dr. Jose Soler, Dr. Sarah Ruepp, Rong Fu, Hao Yu, Jiang Zhang, Anders Rasmussen, Lukasz Brewka, and Ana Rossello Busquet for their help and assistant on reviewing this dissertation. Here I would also like to mention a few researchers at Stanford University whom I have collaborated: Shing-Wa Wong, She-Hwa Yen, Divanilson R. Campelo, Luca Valcarenghi, and Shinji Yamashita. Thanks to all my friends in Denmark who have helped and encouraged me during these years.

It has been a great pleasure to be a member of Networks Technology & Service Platforms group at DTU Fotonik, which brings the special meaning to my education and life. Here I would also like to take this opportunity to express my gratitude to all the staff members of DTU

Fotonik community who has helped in smoothing my transition as a graduate student.

Finally, there seems no word that can express the support and strength I have received from my parents. Thank you, mom and dad, for all I have inherited from you.

Ph.D. Publications

The following publications have been made throughout this Ph.D. project.

Publications on the topic: integrated control platform design in converged optical and wireless networks

[1] Y. Yan, H. Yu, and L. Dittmann, "Wireless channel condition aware scheduling algorithm for hybrid optical/wireless networks," in *3rd International Conference on Access Networks*, 2008

[2] Y. Yan, H. Yu, H. Wang, and L. Dittmann, "Integration of EPON and WiMAX networks: Uplink scheduler design," in *SPIE Symposium on Asia Pacific Optical Communications*, 2008

[3] Y. Yan and L. Dittmann, "Enhanced signaling scheme admission control in the Hybrid Optical Wireless (HOW) networks," in *High-Speed Networks Workshop, co-located with IEEE Infocom conference*, 2009

[4] Y. Yan, H. Yu, H. Wessing, and L. Dittmann, "Integrated resource management for Hybrid Optical Wireless (HOW) networks," in *International Conference on Communications and Networking in China (ChinaCom)*, 2009

[5] S. Wong, Y. Yan, L. Kazovsky, and L. Dittmann, "MPCP assisted power control and performance of cell breathing in integrated

EPON-WiMAX network," in *IEEE Global Communications Conference (GLOBECOM)*, 2009

[6] Y. Yan, L. Dittmann, S. Wong, and L. Kazovsky, "Integrated control platform with load balancing algorithm in hybrid optical wireless networks," in *Third International Conference on New Technologies, Mobility and Security (NTMS)*, 2009

[7] Y. Yan, H. Yu, H. Wessing, and L. Dittmann, "Integrated resource management framework in hybrid optical wireless networks," *IET Optoelectronics Special Issue on Next Generation Optical Access*, 2009

[8] Y. Yan, S. Wong, L. Valcarenghi, S. Yen, D. Campelo, S. Yamashita, L. Kazovsky, and L. Dittmann, "Energy management mechanism for Ethernet Passive Optical Networks (EPONs)," in *IEEE International Conference on Communications (ICC)*, 2010

[9] Y. Yan, L. Dittmann, S. Wong, and L. Kazovsky, "Load balancing in integrated optical wireless networks: Algorithms and evaluation," in *European Wireless (EW2010)*, 2010

[10] Y. Yan and L. Dittmann, "Modeling energy management mechanism in ethernet passive optical networks," in *13th Communications and Networking Simulation Symposium (CNS'10)*, 2010

[11] L. Kazovsky, S. Wong, V. Gudla, P. Afshar, S. Yen, S. Yamashita, and Y. Yan, "Addressing reach extension and reliability weaknesses in next generation optical access networks: A survey," *IET Optoelectronics Special Issue on Next Generation Optical Access*, 2010

Publications on the topic: IPTV traffic management in Carrier Ethernet transport networks

[1] H. Yu, Y. Yan, and M. Berger, "IPTV traffic management in Carrier Ethernet transport networks," in *OPNETWORK 2008*, 2008

[2] H. Yu, Y. Yan, and M. Berger, "IPTV traffic management using topology-based hierarchical scheduling in Carrier Ethernet transport networks," in *International Conference on Communications and Networking in China (ChinaCom)*, 2009

[3] H. Yu, Y. Yan, and M. Berger, "Topology-based hierarchical scheduling using deficit round robin: Flow protection and isolation for triple play service," in *International Conference on Future Information Networks*, 2009

Publications on the topic: QoS provisioning in MPLS/GMPLS networks

[1] H. Wessing, Y. Yan, and M. Berger, "Modelling direct application to network bandwidth provisioning for high demanding research applications," in *5th International Conference on Applied Informatics and Communications (AIC'05)*, 2005

[2] H. Wessing, Y. Yan, and M. Berger, "Simulation bases analysis on dynamic resource provisioning in optical networks using GMPLS technologies," *WSEAS Transaction on Communications*, vol. 5, pp. 9–16, 2006

[3] H. Wessing, Y. Yan, and M. Berger, *Definition of application needs and scale of dynamics in research network infrastructures - additional deliverable on modelling topics.* MUPBED - Multi-Partner European Test Beds for Research Networking Deliverable D2.1a, 2006

[4] Y. Yan, H. Wessing, M. Berger, and L. Dittmann, "Prioritized OSPF-TE mechanism for multimedia applications in MPLS networks," in *5th G/MPLS Networks conference*, 2006

[5] Y. Yan, H. Wessing, and M. Berger, "Bidirectional RSVP-TE for multimedia applications in GMPLS networks," in *10th International Conference on Optical Network Design and Modelling*, 2006

This dissertation only includes work for the topic on integrated control platform design in converged optical and wireless networks.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

With a rapid growth of new services based on multimedia applications, broadband access is demanded due to dramatically increasing in Internet traffic. New service patterns have been propelled from text- and voice-based services to user generated interactive multimedia services. Voice over IP (VoIP), Video conferencing, Video on Demand (VoD), online gaming and High Definition Television (HDTV) broadcasting services will keep growing in the near future and result in increasing bandwidth requirements. In response to this remarkable development, the underlying telecommunication networks have evolved to support new emerging applications with tremendous growth in bandwidth and capacity.

In today's broadband access networks, there has been successful deployment and fast evolution of various new wired and wireless access technologies. Operators have a strong interest to introduce new systems while corporately exploiting their legacy systems. Therefore, scenarios, where an operator is in charge of multiple wired and wireless networks are common. However, simply combining existing optical networks and wireless networks may not provide highly efficient and effective network performances based on their original and static infrastructure. Thus, an update on advanced and collaborative control plane will be introduced to the converged network. There has been growing interest in academia and industry in how the integration of wired and wireless networks should be designed to exploit resources efficiently and to deliver various services to serve both fixed and mobile users.

The viewpoint taken in this dissertation is that achieving optimal performances is even more challenging in integrated optical and wireless networks. This thesis relies on the protocols in current optical and wireless networks and investigates a way to novel integrated control. The aim of this thesis is summarized as:

> **The objective of this work is to present an integrated optical and wireless network architecture and especially focus on protocol design and service provisioning in such highly demanded network environments. The intended performance of a converged network architecture can be improved using a collaborative and cooperative control platform. Novel network protocols and algorithms are proposed and network simulation is carried out for evaluation.**

In this dissertation, both the term of integrated networks and the term of hybrid networks are used to present the convergence architecture of optical networks and wireless networks.

1.1 Background

1.1.1 Evolution of optical access networks

In the wired regime, optical access networks, most notably Passive Optical Networks (PONs), have received a worldwide deployment to provide FTTx (x stands for home, building, curb, etc.), because of the bandwidth enhancement and lower maintenance cost offered by optical fibres. The traditional PON access solution is single wavelength based, known as Time Division Multiplexed PONs (TDM-PONs), where users access and share the bandwidth in time domain. Broadband PON (BPON), Gigabit PON (GPON), Ethernet PON (EPON) have been standardized as TDM-PONs. Next generation PON architectures increase the bandwidth by employing Wavelength Division Multiplexing (WDM) technology. In WDM-PONs, multiple wavelengths are supported over the same fibre on either or both the upstream and downstream directions. Table 1.1 compares standardized EPON and GPON approaches. [20–22]

	EPON	GPON
Standards	IEEE 802.3ah	ITU G.984
Framing	Ethernet	ATM/GEM (GPON Encapsulation Method)
Max. Bandwidth	1 Gbps	2.5 Gbps
Users/PON	up to 32	up to 128
Max. Reach	20 km	up to 60 km

Table 1.1: TDM-PON comparison of EPON and GPON.

1.1.2 Evolution of wireless access networks

In the wireless regime, various wireless access technologies have been developed with an intention to satisfy the continuously growing demand for ubiquitous communication. These different wireless technologies, such as cellular network, Wireless Fidelity (Wi-Fi), World Interoperability for Microwave Access (WiMAX), can locate in the same area and support diverse communication capabilities. Different technologies are adopted for different applications environments and have their own technical specifications, as briefly introduced in Table 1.2. [23–26]

1.1.3 Evolution of converged optical and wireless networks

A convergence of the optical and wireless technologies is proposed for the future broadband access networks. It has created a new design dimension in terms of protocol and architecture design, infrastructure deployment, and control and management. The resultant network will achieve a converged network architecture, where two types of networks complement with each other. Hybrid optical and wireless access networks are proposed in [27–32]. This hybrid architecture 1) provides high-bandwidth and reliable service via the backhaul optical networks, and 2) supports flexible and ubiquitous access connections to the end users via the wireless access networks.

In this work, EPON and WiMAX networks are investigated and modeled in the optical domain and the wireless domain, respectively. EPON has been widely deployed in the access networks. The control system and protocols in EPON have been studied and standardized [33,34].

	3G cellular	**Wi-Fi**	**WiMAX fixed**	**WiMAX mobile**
Standards	3GPP	802.11a/b/g/n	802.16d	802.16e
Operation mode	Centralized mode with base station.	Infrastructure mode with central base station; Ad-Hoc mode with no administrator.	Point-to-Multipoint (PMP) mode; Mesh mode.	
Data rate	14.4 Mbps as peak downlink data rate; 5.8 Mbps as peak uplink data rate.	54 Mbps using 802.11a/g; More than 100 Mbps using 802.11n.	Up to 75 Mbps.	Up to 15 Mbps.
Coverage (typical)	1.6-5 km	<30 m indoor; <300 m outdoor	4-8 km	2-4 km
Mobility	High	Low	No	Middle

Table 1.2: Brief summary of broadband wireless technologies.

On the other hand, WiMAX is a new-generation wireless technique and it has now been standardized. Many network operators and equipment manufactures have expressed their support for WiMAX technology [35]. WiMAX aims to provide broadband support to both fixed and mobile users. The convergence of EPON and WiMAX networks is an attractive alternative for the next generation broadband access system. Therefore, it is of interest to study the interworking problem in the integrated EPON and WiMAX networks.

1.2 Research Contributions

This dissertation makes four important contributions in studying and designing the converged optical wireless networks. These contributions are briefly stated in the following subsections.

- ***Converged optical and wireless networks and integrated control platform***: Chapter 2 first introduces an overview of the optical wireless convergence network architectures. The design of hybrid EPON and WiMAX networks is studied, where the back-end is a wired EPON optical access network, and the front-end is managed by WiMAX wireless access networks. The connection unit in between operates as both the Optical Network Unit (ONU) and the wireless Base Station (BS). An integrated control platform is proposed and the basic idea is to improve cooperation between converged optical and wireless networks. Collaborative control functions, such as uplink scheduling, call admission control, load balancing, and energy efficiency are considered.

- ***Integrated resource management mechanism with call admission control***: In the integrated network, a resource management mechanism is responsible for optimizing both optical and wireless channel utilization. In Chapter 3, a novel call admission control scheme is proposed to examine and qualify wireless connection requests. The expected transmission delay is calculated in the optical domain and disseminated to wireless networks as a parameter in call admission control. It tightly couples the optical bandwidth allocation and wireless call admission control through an enhanced signalling protocol. Therefore, the overall resource utilization is improved with the aid of exchanged information between optical and wireless networks.

- ***Integrated load balancing mechanism***: Chapter 4 proposes and investigates the characteristics of a power control model for the integrated network. The feature of centralized management in EPON allows to assign power levels to multiple wireless networks simultaneously. The cell breathing technology is applied at the wireless network to adjust wireless network coverage sizes and user numbers according to the assigned transmit power. In order to assign proper power levels, an optimization model is formulated and an iterative algorithm is developed with two different scheduling schemes. Two alternative feedback schemes are proposed and compared, which are used to inform the central controller in the optical domain with current wireless network status.

- ***Energy management control in optical access networks***: Chapter 5 explores a major research opportunity in developing an energy management mechanism for the EPON system to pursue power saving. The sleep mode operation is proposed in order to minimize power consumption. An EPON unit turns to low power consumption mode when there are no communication tasks in neither upstream nor downstream directions. Two scheduling schemes are investigated and compared, which are used to assign a sleep period and a wake up period to multiple associated EPON units. The proposed energy management control schemes are modeled. Simulation results show a significant power saving. Furthermore, various network performances of the proposed energy management mechanism are studied.

1.3 Outline of the Dissertation

Chapter 2 provides background knowledge of the EPON and WiMAX technologies and defines a functional overview of the proposed integrated control platform. This work has been published in [1, 7].

Chapter 3 focuses on an integrated resource management mechanism with a call admission control algorithm, aiming to regulate incoming traffic with delay bound. This chapter envisions how the integrated control platform can be deployed to improve cooperation between the original control functions in the optical and wireless network. This work has been published in [2, 3].

Chapter 4 explores a load balancing mechanism, coupled with cell breathing technology, to achieve an optimal solution of power management and traffic distribution for integrated optical and wireless networks. Optimization formulation, intuitive iterative algorithms, and alternative information exchange schemes are designed and presented. This work has been accepted for publication in [5, 6].

Chapter 5 discusses an energy management mechanism for the optical access network. A novel sleeping mode enabled ONU operation is proposed, and two scheduling schemes are proposed to preserve energy consumption and maintain conflict-free communications. The potential moderate degradation of network performances caused by the consideration of power saving as the first design criteria is examined and compared. This work has been published in [8, 10].

Conclusion of this dissertation is given and future research is addressed in Chapter 6.

Chapter 2

Converged Optical Wireless Networks and Integrated Control Platform

2.1 Converged Optical Wireless Network Architectures

One of the most recent strategic issues within the telecommunication industry is the subject of hybrid optical and wireless networks. These networks aim to enable Telecom operators provide a complete environment for deployment, management and administration for a set of wireless and wired technologies. The general context that underlies this dissertation is to exploit the synergy of complementary technologies. Converged optical and wireless networks represent the convergence of cost-efficient fiber-based networks and ubiquitous radio-based networks, where two different and complementary technologies are integrated. Optical communication provides high-bandwidth as backbone network, while wireless communication is deployed as front-end network. It is viewed as an attractive solution and considered as a candidate for the next generation access networks [27–31].

The main driver for the development of the integrated network is the growing demand of high-bandwidth applications. Fiber-based communication systems offer high bandwidth. However, deploying a fiber directly to each home requires high expenses in deployment and maintenance. Moreover, the optical networks have fixed infrastructures and limit coverage. In contrast, wireless techniques provide low cost and mobility to the end-users. Nonetheless, wireless techniques generally operate with low bandwidth. Therefore, the convergence of optical and wireless networks enables the two techniques to complement each other in many aspects.

Recent advances in optical and wireless communications certainly provide ample opportunities of integrated network architectures that benefit the broadband access transmission. Various optical- and wireless-access solutions have been developed over the past decades. In this work, the integrated network is modeled as a convergence of an EPON as a backhaul, and multiple WiMAX networks as the front-end networks, as illustrated in Figure 2.1. In the optical domain, EPON is a cost-effective solution, which uses inexpensive optical splitters to divide the single fiber to individual subscribers and provides easy-to-manage connectivity to Ethernet-based equipment. In the wireless domain, a recent broadband wireless technology, WiMAX, has promised high bandwidth over long range transmission and supports both fixed and mobile transmission. A combination of EPON and WiMAX networks will provide a desirable solution with advantages in terms of low-cost, high-bandwidth and broad-coverage.

2.1.1 Overview of the EPON MAC operations

A typical EPON system consists of one Optical Line Terminal (OLT) functionalized as a central office, one passive optical splitter implemented in the remote node, and multiple Optical Network Units (ONUs) residing at subscribers' locations. IEEE 802.3ah [36] specifies the physical layer and MAC layer characteristics of EPON. The MAC layer operations are briefly introduced in this subsection.

In the downstream direction from the OLT to the associated ONUs, data are broadcasted to each ONU in a point-to-multipoint architecture. On the other hand, it is a multipoint-to-point architecture in the upstream direction from the ONUs to the OLT. The OLT allocates upstream bandwidth among the ONUs and each ONU transmits packets in dedicated time slots. A proper access control mechanism is required

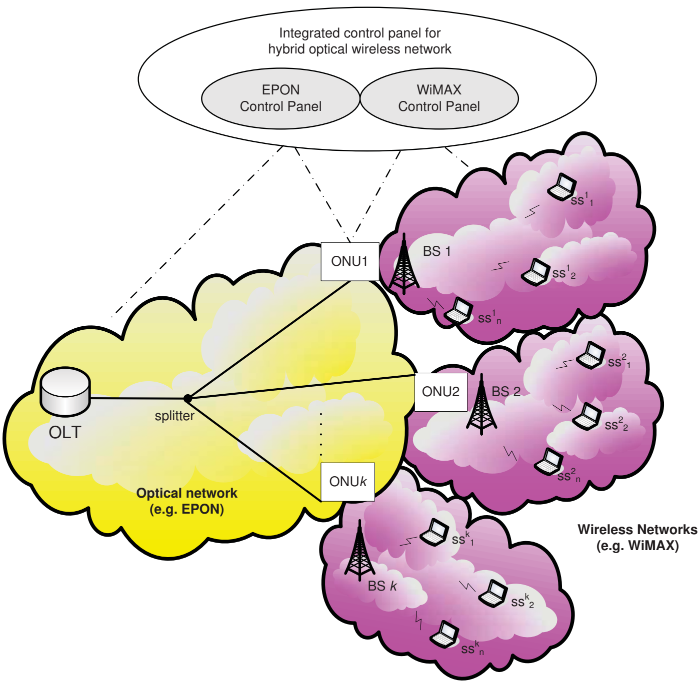

Figure 2.1: Integrated network model with EPON and WiMAX networks.

in order to efficiently coordinate transmission among multiple ONUs.

The signaling access protocol in the EPON MAC layer, known as the MultiPoint Control Protocol (MPCP), is defined in IEEE 802.3ah Ethernet in the First Mile (EFM) Task Force. MPCP provides the signalling infrastructure for coordinating data transmission from the ONUs to the OLT. MPCP uses two types of control messages: the REPORT message is used by an ONU to report bandwidth requirements (normally based on queue occupancies) to the OLT and the GATE message is transmitted by the OLT to issue transmission grants to each ONU.

The upstream bandwidth is divided into subframes in a manner of the Time Division Multiplexing (TDM) and shared by multiple ONUs. The assignment of the limited upstream bandwidth is done by the OLT according to the deployed bandwidth allocation algorithms. Resource management issues in EPONs have been studied in [33,34]. The network resources can be allocated among ONUs either statically or dynamically. In the static approach, the uplink bandwidth is distributed with a fixed amount, which is the Time Division Multiple Access (TDMA) scheme. On the other hand, Dynamic Bandwidth Allocation (DBA) algorithms are proposed to increase the efficiency of bandwidth utilization by adapting to instantaneous bandwidth requirements [37]. Among them, a pioneering and effective DBA algorithm is the Interleaved Polling with Adaptive Cycle Time (IPACT) algorithm [38,39], where its basic idea is that each ONU requests its demanded bandwidth in advance, and the OLT assigns bandwidth to the ONUs based on their requirements.

2.1.2 Overview of the WiMAX MAC operations

WiMAX has evolved from IEEE 802.16a to 802.16e for both fixed and mobile wireless access support. The standards specify the air interface, including the Medium Access Control (MAC) and physical layers. Two modes of MAC operation are specified: Point-to-MultiPoint (PMP) and multipoint-to-multipoint (mesh) [23, 25]. In this dissertation, the PMP operational mode is used.

A centralized Base Station (BS) serves a set of subscriber stations (SSs) and the operation of MAC protocol is connection oriented. Traffic communicated between BS and SS peers are in context of a unidirectional connection and each traffic flow is classified by a Connection

Identifier (CID). SSs medium access is coordinated by the BS through a poll/request/grant mechanism, where bandwidth is requested by an SS and the BS grants uplink bandwidth to an SS. At the beginning of each downlink subframe, the BS broadcasts the uplink and downlink control messages to associated SSs. The control messages are used to notify the SSs of the start time and the end time of the granted transmission periods in the uplink and downlink directions.

In WiMAX, QoS differentiation is defined for various applications, which supports five different traffic types: Unsolicited Grant Service (UGS) supports real-time constant bit rate applications that periodically produce fixed size data packets, Real-time Polling Service (rtPS) is appropriate to real-time applications that periodically generate variable size data packet, extended real-time Polling Service (ertPS) is ensured with a default bandwidth allocation from the BS, non-real-time Polling Service (nrtPS) is appropriate for delay tolerant applications with variable packet size and with a minimum data rate requirement, and Best Effort (BE) services.

There have been many research done for the QoS and the resource management in WiMAX. In [40], the system performance under different traffic scenarios has been evaluated for a WiMAX network working in the PMP mode. To achieve QoS provisioning in WiMAX, the resource management and scheduling algorithms have been discussed in [41–43]. In order to obtain QoS and fairness, wireless channel aware scheduling algorithms are conceived in [44, 45]. In [46], the Call Admission Control (CAC) schemes are investigated in the WiMAX access networks for the provision of QoS guaranteed services. The priority based resource allocation schemes and CAC schemes are proposed and the scheme based on either a Markov decision process approach or a non-cooperative game theoretic approach are addressed in [47, 48], so that the network reward from the real-time application is increased and resource utilization is maximized.

2.1.3 Motivations

The integration of the optical networks and wireless networks offers an attractive and feasible solution to broadband network access. Motivations behind the integration are addressed from aspects of low cost of deployment, high bandwidth provision, and scalable extension of communication coverage.

- **Cost**: The economic limitation prevents fiber from directly reaching individual customers, especially for those areas with low subscriber density. In contract, wireless technique with low deployment costs alleviates such difficulty. By replacing fiber with free radio media, the cost of last-mile transmission is reduced by an integrated optical and wireless network.

- **Bandwidth**: Due to the emergence of Wavelength Division Multiplexing (WDM) technologies, the bandwidth of backbone network has increased substantially. In the integrated optical and wireless network, EPON fills in the gap between the subscriber network and the core network.

- **Coverage**: In the integrated optical and wireless network architecture, high-speed communication is extended by using antenna for wireless distribution. One important good feature in WiMAX technology is its scalability. Without affecting the existing customers, the service provider could install new service areas by adding new base stations as the user demand grows.

2.1.4 Interworking Architectures

Many interoperated architectures have been proposed to support communications in the integrated network, mainly aimed to extend EPON system with wireless transmission penetrating to the local area [27–29, 32]. The interworking architectures can be classified into three categories according to the interdependence between the two access networks.

- ***Independent architectures*** (shown in Figure 2.2a): the EPON and WiMAX are combined as two isolated networks, where the ONU and the BS are implemented in two nodes. There is a common standardized interface of exchanging control and data messages, and there is no direct communication between the ONU and the SSs. This is an easy-to-deploy interworking architecture. The main disadvantages of this independent approach are that:

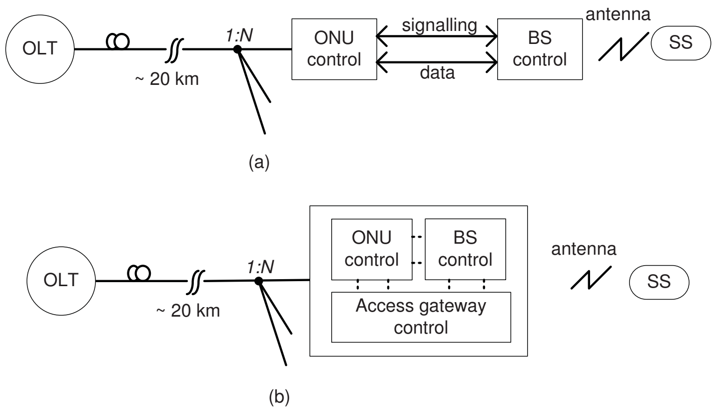

Figure 2.2: Interworking architectures: a. independent architecture; b. Integrated architecture.

  - Since the two domains are separated, control signaling travels between the ONU and the BS, which may introduce relatively high latency.
  - The induced large amount of control and data traffic may cause a network bandwidth bottleneck.

- ***Integrated architectures*** (shown in Figure 2.2b): The ONU and BS functions are implemented and integrated into a single device, where an access gateway control block is designed to coordinate the integration of two domains. The integrated architecture adopts an integrated platform design, which allows to develop an optimal resource management mechanism with significant impact on the overall system. Integrated interworking architecture of optical networks and wireless networks matches well the emerging evolution toward a fixed mobile network infrastructure. With proper control and administration, the optical and wireless domains are seamlessly incorporated through an integrated control platform to provide end-to-end QoS services. The integrated control platform contains cooperative and collaborative control functions located in both the OLT and ONU/BS units.

- ***Radio-over-fiber (RoF) architecture***: RoF based wireless access technology has been researched and proposed as a promising alternative to broadband wireless access network [49]. The RoF architecture consists of a central head-end and a remote antenna unit connected by an optical fiber link, on which the microwave signals are distributed. The advantages of RoF architecture include low attenuation loss, immunity to radio frequency interference and reduced power consumption. However, because RoF involves analogue modulation and analogue signal transmission, the signal impairments, such as noise and distortion, become challenging issues.

In our work, the integrated architecture is deployed, where ONU and BS functions are embedded into a single device, so called Access Gateway (AG). The migration requires changes in the MAC layer and the network layer that bring services to end-users with complied quality of service requirements. In the AG node, a MAC layer mapping and integrating mechanism between the optical domain and the wireless domain needs to be developed, so that is able to translate the communication between two different interfaces. To some extent, the performance of the hybrid network solution relies on the design of AG nodes, which plays an important role for the EPON and WiMAX integration.

2.2 Design of an Integrated Control Platform

The optical and wireless access technologies are originally designed to address different issues and deployed in different scenarios. Therefore any simple combination of them cannot derive optimal network performances. The two parts of the integrated network, the optical domain and the wireless domain, have to be considered and designed altogether. That is, the network conditions, in terms of link resource usage and traffic load, are monitored and exchanged under both the optical and wireless networks. The heterogeneous control plane of the integrated network becomes a key design issue, which can be exploited to maximize the overall resource utilization and optimize the overall network QoS performance.

Traditionally the mechanisms of resource management are designed separately and exclusively for either optical networks or wireless networks. The characteristics and challenges of the hybrid network are ignored. Thus, this has led to a need for an integrated resource management for integrated networks. There are two approaches to manage both optical and wireless channel resources simultaneously for the overall integrated architecture.

- The first approach is to design a resource management scheme from scratch, which deploys newly designed signaling, resource allocation, call admission control, and scheduling algorithms. This approach often incurs a design of new network architectures.

- The second approach is to investigate the existing resource management methods and implement them with modifications and enhancements in the integrated network. This approach is much more agile, effective and has less time-to-market than the previous one. In a practical network development, these two approaches can complement each other. In this work, the second approach is chosen and studied.

Figure 2.3 depicts an overview of functions proposed in the integrated control platform. Basic functions in the EPON and WiMAX MAC layers are modified and extended to support cooperative signalling access between the optical and wireless domains. The whole integrated system is controlled by a set of protocols that manage both wired and wireless networks. In this study, the following control functions are considered:

- ***Call Admission Control (CAC)***: Call admission control is a provisioning process to limit the number of call connections into the network in order to reduce the network congestion and call dropping. In integrated networks, another regulation is added: call connection failure is possible due to violation of QoS requirements in one of two adjacent networks. Although intensively studied in the latest few years for wireless networks, CAC becomes more complicated in the hybrid network architecture. The AG receives numerous requests to set up upstream connections across optical and wireless networks. Traffic flows experience delays from both the optical and wireless domains. The CAC algorithm considers the end-to-end delay across the WiMAX and EPON network, for each connection request to guarantees QoS. For the reason that

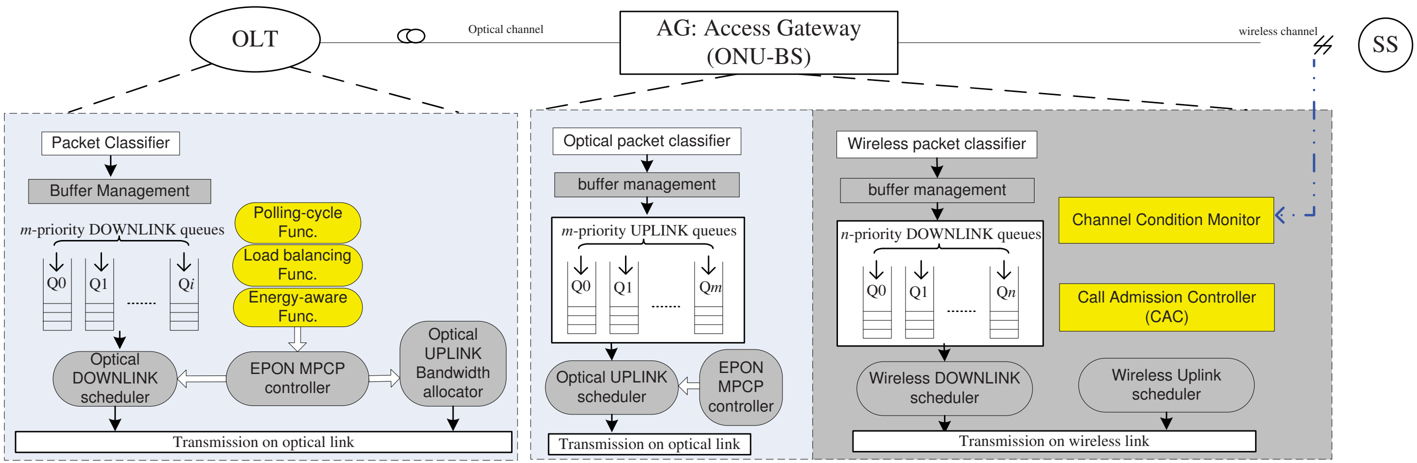

Figure 2.3: Overview of functions proposed for the integrated control platform.

without considering the queuing delay and polling delay in the optical system, the base station in the wireless domain may grant bandwidth to non-conforming traffic, for which QoS requirements (e.g. delay) cannot be satisfied. In this work, an integrated admission control is achieved by the cooperation of the CAC function at the AG and the EPON MPCP controller at the OLT (shown in Figure 2.3). The CAC decides whether to accept connection requests or reject unserviceable requests based on the potential delay calculated by the polling cycle function, which is implemented in the EPON MPCP controller at the OLT.

- ***Packet Scheduling***: Packet scheduling is an important function in the MAC layer to achieve fairness and maximization of channel utilization. Intensive research of upstream scheduling algorithms for EPON system can been found in the literature. In the hybrid optical wireless network, cross-domain traffic flows arrive at the gateway node from mobile users via wireless upstream links, which are subject to signal attenuation, fading, interference and noise. Thus, these existing algorithms cannot be directly adapted without taking features of wireless links into consideration. For the integrated network, a channel-aware scheduling algorithm should be considered, which assigns the optical uplink bandwidth to connections according to the corresponding radio channel quality in the wireless domain. The optical uplink scheduler allocates the assigned optical uplink bandwidth to the multiple queued wireless connections. In a conventional schedulers design in an EPON system, the packet and queue information, such as traffic types and queue sizes, is measured as priority metrics. Designing an integrated scheduling algorithm, the optical uplink scheduler function communicates with the channel condition monitor function to obtain channel conditions. When the quality of a channel drops below a threshold value, no traffic can be transmitted. Once a connection is detected suffering bad channel condition, the future incoming data rate could be predicted to decrease. In this work, an integrated upstream packet scheduler is achieved by the cooperation of the channel condition monitor and the optical uplink scheduler at the access gateway (shown in Figure 2.3). The basic idea is to increase the throughput of connections in good channel condition. The advantage of the channel-condition aware scheduling algorithm compared to priority based one is the QoS assurance of low priority classes with "good" channel condition. That is, the excessive bandwidth of a lightly loaded class is shifted to the highly loaded one to meet the urgent bandwidth demand.

- ***Load Balancing***: In a broadband wireless access network, the base station power is typically managed to minimize mutual interference among adjacent cells. By lowering the transmit power, a base station could further reduce the coverage dynamically during heavy network load. Reducing the coverage could reduce the number of subscribed users and relieve itself from the excess traffic burden. Similarly, a neighboring base station could share the excess traffic load by expanding its coverage. This operation is called cell breathing, and such dynamically balanced network could provide better QoS within a single cell or statistically support more mobile users. At any one time, unbalanced load could occur in wireless networks either due to over-utilization or due to poor channel conditions. To determine whether cell breathing is necessary, a central controller calculates the expected total service time and equalizes them across all wireless networks. During cell breathing, wireless networks could loose connections to some subscribers. In this work, an integrated admission control is achieved by the cooperation of the channel condition monitor at the AG and the EPON MPCP controller at the OLT (shown in Figure 2.3). The OLT makes joint processing of traffic and cell power assignments based on the current wireless network status in terms of backlog data size and channel conditions.

- ***Energy Efficiency Control***: In energy-efficient network systems, the sleep mode operation is introduced. Nodes (or stations) are allowed to switch to the sleep status when they are idle and to wake up when they receive or transmit packets. There is a trade-off decision between maximizing the power saving and guaranteeing the network performance at the same time. The design of an efficient power saving control scheme needs to be a feasible solution to both enable power saving and minimize the impact of traffic. The main idea is to put ONUs into the sleep mode and determine a suitable wake-up time schedule at the OLT by means of jointly scheduling upstream and downstream transmission periods. In this work, the energy efficiency is achieved by introducing a new function at the OLT (shown in Figure 2.3). The energy consumption for an EPON model is examined and evaluated. Two downstream scheduling algorithms are analyzed to enable the power saving mode. Necessary extensions are introduced to the traditional MPCP control messages and formulate the energy consumption equations in both downstream scheduling cases.

2.3 Summary

In this chapter, the architecture and a vision for the integrated optical wireless networks are introduced. Motivations are presented to articulate the trend of combining the complementary characteristics of optical and wireless technologies. The main focus of this work is placed on the design of an integrated control platform in order to maximize network performances. Among cooperative control functions, the admission control, load balancing, packet scheduling, and energy efficiency functions of the integrated control platform are discussed.

Chapter 3

Enhanced Signalling Scheme with Admission Control in the Integrated Control Platform

The hybrid optical wireless network has been presented as a promising solution to meet the increasing user bandwidth and mobility demands. Due to the basic differences in the optical and wireless technologies, a challenging problem lies in the Media Access Control (MAC) protocol design, so that it can support stringent Quality of Service (QoS) requirements. Efficient utilization of available bandwidth over hybrid optical wireless networks is a critical issue, especially for multimedia applications with high data rates and stringent delay requirements. In this chapter, a resource management framework is presented for a hybrid EPON and WiMAX network. At first, the problem of sharing resource allocation information between optical and wireless domains is formulated and studied. The existing signaling protocol in EPON is modified and enhanced. Secondly, a novel integrated admission control scheme is proposed. Simulation results show the performances of the system in terms of throughput, delay and packet dropping probability under different system parameters. These parameters include the frame duration, the traffic load and the total number of shared users. The results also highlight that the proposed system achieves significant improvements over the traditional approach in terms of user QoS guarantee and network resource utilization.

3.1 Introduction

Recent years have witnessed the explosive growth of multimedia applications, which are characterized by high speed and stringent QoS requirements. An integration of optical and wireless technologies has been viewed as a viable solution to delivery the quadruple play services (data, voice, video and mobility). A hybrid EPON and WiMAX network is an example of the integrated network, where EPON deployed as the backhaul network and multiple WiMAX networks connected as the front-end networks.

In order to design an unified control system in the hybrid network, an integrated resource management needs to be considered and accomplished. It can dynamically allocate network resources based on the current wireless network and optical network conditions. The objective is to improve the existing protocols and to design an integrated resource management in the integrated EPON and WiMAX network. Moreover, it can optimally admit the requests and allocate the network resources in order to maximize the resource utilization and satisfy the prescribed QoS requirements.

In the integrated optical wireless network, resource management is implemented in a centralized architecture, where an intelligent control mechanism in the EPON polls and allocates bandwidth to all access gateways, which connect with wireless systems. There are several methods that can be used to achieve efficient resource management, such as using dynamic bandwidth allocation algorithms and admission control schemes. The resource management functions should be designed with consideration for QoS provision, such as packet delay, network throughput and so forth. In the hybrid optical wireless network, the resource management is composed of two stages. The first stage is the optical resource allocation among attached access gateways in the optical domain. The second one is the radio resource allocation among connected mobile users in the wireless domain. Appropriate cooperation and integration of resource allocations in these two domains are crucial to investigate.

The node located at the common boundary of these two domains needs an advanced resource management, in order to coordinate the communication between the optical and the wireless domains. In this chapter, QoS performance enhancements by determining proper admission decision over the whole integrated network are presented.

The framework consists of a modified resource negotiation scheme in EPON and a polling-delay-aware admission control scheme in WiMAX. The admission control scheme is located at the interface node of the optical and wireless networks and collects network information from both networks. Thus, the potential network congestion is able to be detected and avoided. The achievements of this work include:

- calculation of the joint delay time while packets transmit through the wireless and optical domain;
- guarantees of bandwidth allocation for the requests with conformed traffic;
- maximization of network throughput from QoS guaranteed flows.

The rest of this chapter is organized as follows. For completeness as background knowledge, Section 3.2 briefly introduces related work in the integrated resource management scheme . In Section 3.3, following a description of the hybrid system model, a two-step resource management approach is proposed. The proposed integrated resource management mechanism with an enhanced MPCP scheme and an integrated admission control scheme is addressed in detail in Section 3.4 and Section 3.5, respectively. Simulation results of the proposed approach are provided and discussed in Section 3.6. Conclusions are given in Section 3.7.

3.2 Related Work

In order to achieve a highly efficient integrated resource management mechanism, several aspects should be addressed in advance. The main technical issues are: (1) routing algorithm, (2) packet classification and scheduling mechanism and (3) resource management and call admission control mechanism.

At first, a routing algorithms is used to locate the optimum routes for packet transmissions in the hybrid network. In [50], an integrated routing algorithm is proposed to select the optimum routes for transmissions in both optical and wireless domain. The routes are computed in the optical domain, while the decision is based on the exchanged information from wireless domain, such as wireless link state and average traffic rate. For the front-end wireless network in integrated networks, a delay-aware routing algorithm is proposed in [51], which selects a path with the minimum predicted delay.

Second, traffic from various applications should be treated differently. Since the granularity of classifying different services (DiffServ) are defined distinctly in the optical network and the wireless network, a unified packet classification scheme is required for the integrated network. Packet scheduling algorithms have been investigated extensively for the wired [52] and wireless network [53] over the past decades. Packet scheduling algorithms are used to select the sequence of transmitting packets, while satisfying QoS requirements. In the reference [54], capacity analysis and wavelength multiplexing are exploited, and a centralized MAC layer is designed with discussions on scheduling and resource sharing policy. In work [1], a priority-based scheduler is proposed for the uplink stream transmission in the integrated network. Its key feature is that the scheduler prioritizes connections in the optical domain, while taking wireless channel quality into account.

At last, for the resource management and CAC issues, an integrated resource management is required to ensure the QoS over the overall integrated network. Bandwidth is allocated and distributed according to the applied bandwidth allocation schemes. The reference [55] presents an integrated wireless and SONET architecture. An optimal utility based bandwidth allocation scheme is designed for multimedia application. The reference [56] investigates a dynamic bandwidth allocation (DBA), which takes into consideration the specific features of the converged network to enable a smooth data transmission across optical and wireless networks.

Research work on CAC for wireless technologies, such cellular networks and WLAN, has been discussed in [46, 57, 58]. For integrated optical and wireless networks, as proposed in [59], a central controller is designed in the optical domain to maintain an integrated view of the hybrid network. The connection requests submitted in the wireless domain are delivered to the optical domain. Moreover, a QoS-aware scheduling scheme is used in the call admission control scheme, which aims to increase network throughput while reducing overhead of the cross-domain (between the optical domain and the wireless domain) communication. In this chapter, CAC is achieved with an enhanced signaling protocol and admission control scheme, which has not been discussed previously in the integrated network.

3.3 System Model and Problem Definition

An integrated EPON and WiMAX network consisting of two parts, the backhaul EPON infrastructure and the front-end WiMAX networks, is considered as the system model. The node with integrated ONU and BS functions is referred to as an ***Access Gateway*** (***AG***). As shown in Figure 3.1, there are K AGs at the interface and there are N Subscriber Stations (SSs) connected in the WiMAX network. In this integrated architecture, the upstream traffic is first aggregated at an AG and then forwarded to the OLT. The upstream transmissions from the associated ONUs are scheduled in a TDM manner. Each ONU transmits during the assigned transmission period in order to avoid a traffic congestion at the splitter. For the downstream traffic, packets are first transmitted to the AG and then forwarded to each SS in the allocated channel slots.

The AG is the interface between the EPON and WiMAX networks. It handles communications as within its own wireless network (e.g. a single-domain connection), and across both optical and wireless networks (e.g. a multi-domain connection). QoS improvements for the single-domain connections in the WiMAX network has been discussed extensively. This work focuses on providing QoS guarantees to the multi-domain connections, because the new features of an integrated network are highlighted in the multi-domain communication.

To set up a multi-domain connection in the hybrid network (a connection from an SS to the OLT), the SS initiates a request to the corresponding AG. The acceptance or rejection of a request is determined by the admission control scheme in the AG. A request can be accepted if there is enough buffer size and the QoS requirements can be satisfied. A channel resource control is assumed to manage the optical channel and wireless spectrum allocation and scheduling in the integrated network. It can collect information, including traffic demand and channel avail-

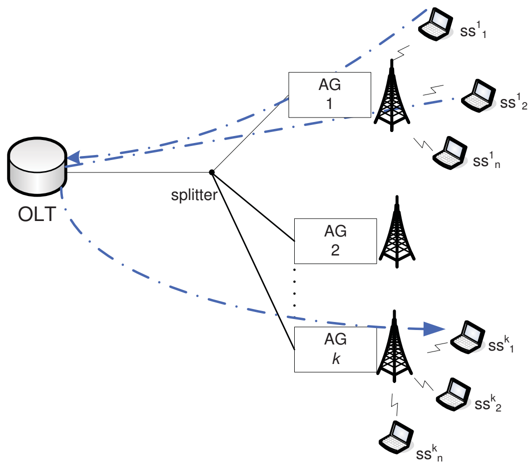

Figure 3.1: Illustration of the integrated optical wireless network architecture

ability, from both optical networks and wireless networks. The central controller computes the scheduling and channel allocation solution for a multi-domain connection. The considered performances are the QoS experiences for both optical and wireless domains.

As mentioned in the previous section, the bandwidth allocation and admission control algorithms have been extensively researched in EPON and WiMAX networks, respectively. However, these are all studies in either EPON or WiMAX, and no model exists to evaluate the network performances in the case of transmissions through both the optical and the wireless domain. In the hybrid optical wireless network, numerous requests are transmitted to the AG to set up upstream connections from an SS to the OLT. Admitting a new connection request in a WiMAX network, without considering the queuing delay and polling delay in the following EPON system will grant bandwidth to nonconforming traffic, e.g. to grant bandwidth to a service request, for which the delay requirement cannot be met. Allocating transmission opportunities to nonconforming traffic will cause waste of scarce bandwidth and QoS degradations to already conforming traffic.

In order to provide a sustainable and consistent hybrid access network, an integrated resource management framework on the upstream channel is essential. The basic idea in the enhanced admission control mechanism is to integrate information derived from both optical and wireless domain and then take the overall delay into account to provide optimal admission decisions. In order to calculate the overall delay, the network information needs to be exchanged between the optical and wireless networks. Our proposed approach consists of two schemes: (1) an enhanced signalling protocol and (2) an integrated CAC scheme. In the following sections, the existing control functions are investigated. Furthermore, enhancement and improvement are presented in details.

3.4 Integrated Resource Management Framework Part I - Enhanced MPCP Scheme

3.4.1 Basic MPCP Operations

As introduced in 2.1.1, two control messages, REPORT and GATE, are defined in the MPCP as signaling between the OLT and AGs at the EPON domain in the hybrid network. The GATE message is initiated by the OLT and sent to AGs carrying granted bandwidth information. The REPORT message is directed to the OLT and used to indicate the residual queued data size in AGs.

Two processes are performed: the discovery process and the normal process. During the discovery process, the OLT searches for AGs, registers attached AGs, and calculates the the Round Trip Time (RTT). The RTT is calculated at the OLT using the stamped packet and the value is determined by the distance between the OLT and the AG. The discovery process is repeated periodically, so that newly connected AGs can seamlessly be added without interrupting the current network operation. After discovering and registering the connected AGs, the OLT sets up an entry table, which contains the AG Logic Link Identification (LLID) and the round trip time (T_{rtt}).

A key perspective of the normal process is the ability to assign bandwidth and schedule transmission for all registered AGs in a manner of fairness and without conflict. In the upstream direction, multiple AGs share a common optical medium. The upstream transmission period is divided into multiple time slots. Each ONU can only transmit during its own slot time. The OLT polls registered AGs and assigns time slots either statically based on the TDMA scheme or dynamically based on the resource requirement negotiation, e.g. IPACT.

- **TDMA scheme**: each AG is granted a fixed time slot length for upstream transmission. TDMA scheme is simple to implement. However, without considering the instantaneous changes of traffic, the OLT assigns some time slots to AGs under very light load. This leads to the bandwidth resources being under-utilized.

- **IPACT scheme**: a resource negotiation is applied between the OLT and the AGs. The OLT polls AGs and issues transmission grants to them in a round-robin fashion. The AG generates the request message and reports the queued data size to the OLT. The request data size is used to determine the grant transmission period for the next grant of the AG. Therefore, bandwidth is dynamically assigned to AGs according to their queue occupancies. The OLT keeps track of the T_{rtt} of all AGs . The value of T_{rtt} is computed based on the distance between the OLT and the attached AG. By being aware of the round trip time, the OLT can calculate the arrival time of the REPORT message from each AG. As a result, the OLT can grant the ONUs and schedule their upstream transmission. By dynamically allocating time slots based on instantaneous buffer size, the bandwidth utilization is improved.

In this thesis, a ***polling cycle*** (T_{cycle}) refers to a period in which all AGs are served and a ***granted subframe period*** (T_{sp}) refers to a period, which is assigned to an AG for the uplink transmission. The order of transmitted polling messages is determined by the deployed scheduling algorithm (e.g. simple round robin fashion or priority based). In this work, the round robin scheduling scheme is considered.

3.4.2 Enhanced MPCP Operations

In the integrated architecture, an AG performs resource management operations for both optical and wireless networks. In order to calculate the packet transmission delay in the optical domain, the AG needs to know the estimated waiting time for its next poll. This extra information can be derived from the OLT using a modified GATE message. Originally defined in MPCP, the GATE message consists of a 1-byte granted start time (t_{start}) and a 2-byte granted bandwidth. There is a new 2-byte field, T_{next}, is added to original MPCP GATE control message. The field T_{next} indicates the interval between the time of the current and the next polling operations to an AG. In this section, t is used to indicate the instant time and T is used to indicate the duration between two time points (e.g. $t_x - t_y$).

After the OLT computes the start time and subframe length for each AG, the OLT can obtain the total transmission length (T_{cycle}) for all k AGs (as illustrated in Equation 3.1).

$$T_{cycle} = \sum_{i=1}^{K} \left(\frac{BW^i}{R_o} + T_g \right) \tag{3.1}$$

where BW^i is defined as the granted bandwidth as the size of transmission capacity (in bits). R_o is the transmission rate of the optical uplink and T_g is the guard time between two successive upstream transmissions. After a period of T_{cycle}, an AG is polled again. The interval between two adjacent polling operations, the next cycle time, can be calculated and added into the original GATE message as a new field. The GATE control message can be embedded within an Ethernet frame. The receiver recognizes the GATE control message and then extracts the following 6-byte grant information. The modified GATE message can be used when either the TDMA or the IPACT bandwidth allocation scheme is deployed in the optical domain in the integrated networks.

3.4.3 Enhanced MPCP Operations under TDMA scheme

The modified resource sharing scheme under the TDMA scheme is shown in Figure 3.2. The polling sequence of AGs is scheduled in the OLT. Here a round robin scheduler is assumed so that all AGs are polled in turn. We assume the scheduling principle will not change. The value of granted bandwidth (BW^i) is fixed and same to all AGs. In other words, the values of polling cycle (T_{cycle}) are same for each AG. The expected next cycle time for the i^{th} AG is calculated in Equation 3.2. The value

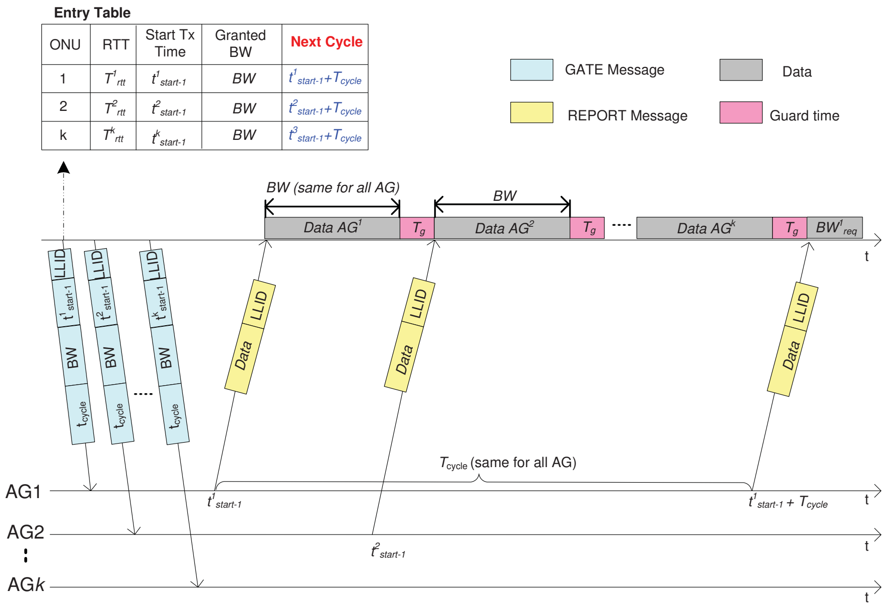

Figure 3.2: Illustration of the modified GATE message with a new field, the next cycle time, is used with TDMA scheme.

of t_{next} for each AG is computed and transmitted to the AG via the GATE message at the beginning of each cycle.

$$\begin{aligned} t^i_{next-TDMA} &= t^i_{start} + T_{cycle-TDMA} \\ &= t^i_{start} + K \cdot \left(\frac{BW^i}{R_o} + T_g \right) \end{aligned} \tag{3.2}$$

3.4.4 Enhanced MPCP Operations under IPACT scheme

Compared to the TDMA scheme, the IPACT scheme is more complicated. The AGs are polled by the OLT in an interleaved manner. The $(i+1)^{th}$ AG is polled via the downlink GATE message, while the i^{th} AG is in the uplink transmission. The distribution of upstream bandwidth is based on AG's requests rather than on a fixed generic amount. The modified resource sharing scheme under the IPACT scheme is shown in Figure 3.3.

After the discovery process, the OLT fills the entry table with initial information (e.g. node ID, RTT value and granted bandwidth) of AGs. The granted bandwidth, BW^i, is assigned for the i^{th} AG as the scheduled upstream transmission period. When the discovery process is completed, the OLT starts to poll the AGs via the GATE message. Here a round robin scheduler is assumed so that all AGs are polled in turn. As explained in [37], the OLT grants AGs in an interleaved manner and distributes the length of the time slot based on their requests. For each AG, the polling time and the start time for its upstream transmission are computed using Equation 3.3 and Equation 3.4. As illustrated in Figure 3.3, t^i_{poll} is the time when the OLT polls the i^{th} AG. At this time, the GATE message is transmitted with the bandwidth grant information. Therefore, after receiving an updated bandwidth request from the i^{th} AG, the time to start upstream transmission in the $(i+1)^{th}$ AGs calculated in Equation 3.4.

$$t^{i+1}_{poll} = t^i_{poll} + T^i_{rtt} + \frac{BW^i}{R_o} + T_g - T^{i+1}_{rtt} \tag{3.3}$$

$$\begin{aligned} t^{i+1}_{start-IPACT} &= t^{i+1}_{poll} + \frac{T^{i+1}_{rtt}}{2} \\ &= t^i_{poll} + T^i_{rtt} + \frac{BW^i}{R_o} + T_g - \frac{T^{i+1}_{rtt}}{2} \end{aligned} \tag{3.4}$$

When the IPACT scheme is deployed, the period of a polling cycle is determined by the total granted bandwidth (as illustrated in Equation 3.5). The value of $T_{cycle-IPACT}$ needs to be calculated, when there is new update on the bandwidth assignment. as illustrated in Figure 3.3, after AG^1 is polled, AG^1 transmits REPORT message, including LLID, required bandwidth (BW^1_{req}) and data. The OLT updates the entry table while the request bandwidth is granted (the BW^1_1 is replaced by BW^1_2, where BW^i_j means the bandwidth for i^{th} AG in the j^{th} round of polling). For example, the $T_{cycle-IPACT}$ for AG^2 is computed, after the OLT updates the entry table with new granted bandwidth (BW^1_2).

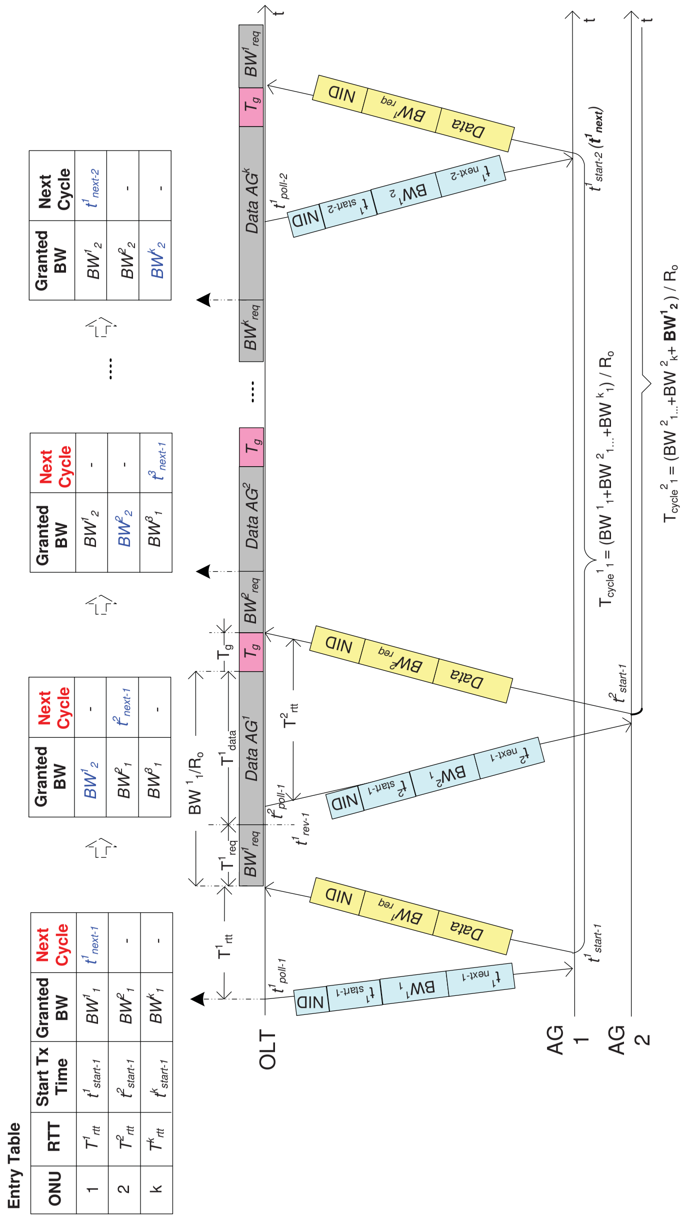

Figure 3.3: Illustration of the modified GATE message with a new field, the next cycle time, is used with IPACT scheme.

$$T^i_{cycle-IPACT} = \sum_{i=1}^{K} \frac{BW^i}{R_o} + K \cdot T_g \tag{3.5}$$

The new field, $t_{next-IPACT}$, is computed and sent to the next polled AG using Equation 3.6. The value of $t_{next-IPACT}$ indicates the waiting period for an AG to its next poll, which is calculated according to the total number of connected AGs and their assigned bandwidth.

$$t^i_{next-IPACT} = t^i_{start} + T^i_{cycle-IPACT}, \quad i \in K \tag{3.6}$$

For the reason that granted bandwidth is changed according to i^{th} AG's request, the value of $T_{cycle-IPACT}$ becomes different in cycles. The entry table is updated and the $t_{next-IPACT}$ is computed for the $(i+1)^{th}$ AG when the REPORT message from i^{th} AG is received. It is noticed that the request bandwidth from the i^{th} AG need to be received before the OLT sends a grant message to the $(i+1)^{th}$ AG in order to calculate $t_{next-IPACT}$ for the $(i+1)^{th}$ AG. In another words, the time to update the entry table (t^1_{rev-1}) should before the time to poll the next AG (t^2_{poll-1}). This condition is expressed using Equation 3.7.

$$\begin{aligned}
& t^{i+1}_{poll} > t^i_{rev} \\
& \Rightarrow \quad t^{i+1}_{poll} > t^i_{poll} + T^i_{rtt} + T^i_{req} \\
& \Rightarrow \quad t^i_{poll} + T^i_{rtt} + BW^i/R_o + T_g - T^{i+1}_{rtt} > t^i_{poll} + T^i_{rtt} + T^i_{req} \\
& \Rightarrow \quad t^i_{poll} + T^i_{rtt} + (T^i_{req} + T^i_{data}) + T_g - T^{i+1}_{rtt} > t^i_{poll} + T^i_{rtt} + T^i_{req} \\
& \Rightarrow \quad T^i_{data} + T_g > T^{i+1}_{rtt}
\end{aligned} \tag{3.7}$$

In above equation, the assigned upstream bandwidth includes 1) the transmission of a REPORT message (T^i_{req}), which is used to carry the bandwidth request information; 2) the transmission of payload (T^i_{data}). It is shown that the entry table can be updated and the polling cycle delay can be calculated if the transmission period of payload for last polled AG (T^i_{data}) plus the guard time (T_g) is larger than the return trip time (T_{rtt}^{i+1}). Under these circumstances, the GATE message to the next AG is able to be transmitted after the arrival of the REPORT message from the i^{th} AG. When this condition cannot be satisfied, the polling cycle time for next AG is calculated with a maximum grant bandwidth, in stead of the actual request bandwidth value.

3.5 Integrated Resource Management Mechanism Part II - Integrated Admission Control

An uplink transmission scenario from an SS to the OLT through the integrated optical wireless network is considered. Although EPON and WiMAX use different channel access mechanisms, they both need admission control to determine how much traffic can be handled in the optical and the wireless domains separately, so that the prescribed QoS for each traffic stream can be maintained. Implementing two individual admission control schemes for each network is not effective and efficient. An Integrated Optical Wireless Admission Control (IOW-AC) scheme is proposed to cope with this problem, which provides delay bounds to multidomain connections. This chapter first presents a model to evaluate the delays experienced by the requests from the SSs, and then describes the proposed IOW-AC scheme.

3.5.1 System description

As illustrated in Figure 3.4, AGs transmit traffic to the OLT on the uplink channel within their respective assigned slot times. The assigned transmission period can be equal and fixed to all AGs or be different and dynamic. If it is dynamically changed based on the AG's request, the GATE and REPORT messages are exchanged between the ONU and the OLT before the upstream transmission starts.

Since the IEEE 802.16 MAC protocol is connection oriented, an SS first initiates a connection with an AG and informs the associated service flow (i.e., real-time traffic or BE traffic). The SS request, with the QoS requirements from the application layer, is passed to the admission control mechanism and the AG grants the request based on pre-defined criteria, such as the queue size and the delay. The connection signalling

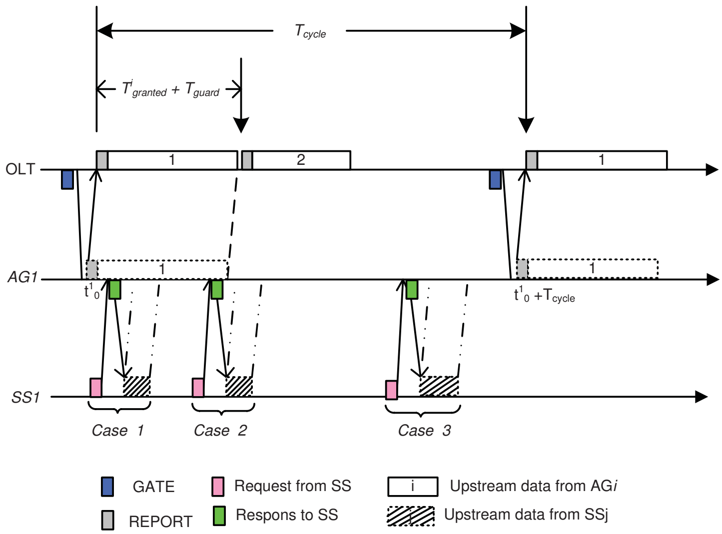

Figure 3.4: The upstream communication in the hybrid optical wireless network.

messages (request and response) between the SS and AG are defined in IEEE 802.16 standard. Once the connection is set up, the packets received at AG are classified and forwarded to the appropriate queue.

The IOW-AC is aware of the transmission status in the optical domain. Upon receiving a GATE message from the OLT, the AC receives knowledge about the length of the optical uplink transmission period ($T_{granted}$) and the length of waiting time to the next polling cycle (T_{cycle}). The $T_{granted}$ consists of T_{req} and T_{data}. The selection of the queued packets into the upstream $T_{granted}$ is determined by the AG based on the deployed scheduling algorithm, e.g. evenly or proportionally distributed among different service queues. In this chapter, a strict priority based scheduling algorithm is adopted. At the beginning of an uplink transmission to the OLT, t_0, a scheduler is used in the AG to select packets from queues of different service classes. Every service class may be guaranteed a minimum bandwidth (V_{min}) and a maximum bandwidth (V_{max}) so that all service classes receive their shares of the scheduled uplink channel. The scheduler selects packets from each service queue based on the priority of traffic types, where the real-time traffic has the highest priority and the BE traffic has the lowest. The number of packets is proportional to the incoming throughput of each traffic class.

Based on the traffic type, the queue of AG is logically divided into several subqueues, e.g., voice traffic is sorted into the $Q_{sub_{voice}}$ subqueue, video traffic is sorted into the $Q_{sub_{video}}$ subqueue, and BE traffic is sorted into the $Q_{sub_{BE}}$ subqueue. Within a granted time slot, the bandwidth is distributed to each type of traffic with a maximum limit, e.g. $V_{max-voice}$, $V_{max-video}$ and V_{max-BE}.

3.5.2 Delay analysis

As explained in Section 3.4, the i^{th} AG receives the estimated next transmission time t^i_{next} from the OLT via the GATE message. On the other hand, the AG receives requests from its connected SSs. Assuming that the traffic arrival time, t, of any requests from an SS_j to the corresponding AG^i are uniformly distributed in the interval $[t^i_0,\, t^i_0 + T_{cycle}]$, where t^i_0 indicates the starting time of the subframe AG^i. The probability of an SS traffic arriving at the i^{th} AG during the time of an uplink transmission to the OLT ($T^i_{granted}$) will be Equation 3.8:

$$\boldsymbol{P}[t^i_0 \leq t < (t^i_0 + T^i_{granted})] = \frac{T^i_{granted}}{T^i_{cycle}} \tag{3.8}$$

When traffic arrives at the AG^i from the wireless domain, AG^i is able to estimate the waiting time for the traffic to be served. Calculated as in Equation 3.9, the overall estimated delay (d_{est}) includes the waiting time until AG^i is polled by the OLT ($d_{polling}$), the waiting time for the prior data in the same queue being served ($d_{queueing}$) and the wireless transmission ($d_{tx-wireless}$) and propagation delay ($d_{prop-wireless}$).

$$d_{est} = d_{polling} + d_{queueing} + d_{tx-wireless} + d_{prop-wireless} \tag{3.9}$$

Where $d_{polling}$ is calculated based on the next cycle time (t_{next}) and the packet arrival time (t). The t_{next} is computed as explained earlier. $d_{queueing}$ is determined by the current subqueue size and the assigned bandwidth within a subframe. If the subframe has enough empty space, the new coming packet can be served right after the queued packets in the following subframe. Since the queuing delay is not a focus of this work, it is assumed that the available buffer size in AG is enough. $d_{tx-wireless}$ is calculated based on the packet size and the wireless link rate.

$d_{polling}$ is determined by the information provided in the GATE message (e.g. the starting transmission time t_0, the granted slot time $t_{granted}$ and the next transmission time t_{next}). As shown in Figure 3.4, there are three different cases, where AG receives the connection request from SSs and the polling delay is calculated.

- Case1: the request of j^{th} SS, SS_j, arrives when the AG^i is being polled. A request can be accepted and served within the current subframe only if the following two conditions are fulfilled: (1) the corresponding packets can be received before the current subframe is finished; (2) there is enough space in the ongoing upstream subframe to carry the arriving packet from the SS at the instant t. In this situation, Case1a, the polling delay is $d_{polling} = 0$. Otherwise, as Case1b, the arriving packet will be stored in the AG's buffer and transmitted in the next assigned upstream subframe. Therefore the polling delay is $d_{polling} = t_{next}$ - t.

- Case2: the request of SS_j arrives when the AG^i is being polled. If the corresponding packets can not be received (at time t) before the current subframe is finished, the request has to wait until the next cycle to be served. Thus, $d_{polling} = t_{next}$ - t.

- Case3: the request of SS_j arrives when the AG^i is waiting to be polled. The earliest time for the request to get served is the next polling time. Thus, $d_{polling} = t_{next}$ - t.

In the case $t \in [t_0^i, t_0^i + T_{granted}^i)$, the request arrives at the time when the AG is being polled. The polling delay is determined by the packet arrival time and the available amount of buffer. If the request of SS_j arrives at the AG^i, which is waiting to be polled, i.e., $t \in [t_0^i + T_{granted}^i, t_0^i + T_{next-tx}^i)$, the packet has to wait until the remaining part of an uplink transmission from other AGs to the OLT is completed. It is noted that an increase on the subframe size $T_{granted}^i$ can result in a larger number of accepted requests, but it also increases the polling delay.

Thus, the overall estimated delay for a connection request from SS to the OLT is calculated as indicated in Equation 3.10. L_{packet} is the length of upstream packets of SS_j. $R_{wireless}$ is the wireless transmission rate of SS_j. $T^i_{j,prop-wireless}$ is the propagation time from SS_j to AG^i; t^i_0 is the starting time of the subframe of AG^i and t^i_{next} is the time when next polling message arrives.

$$d_{est} = \begin{cases} \frac{Q^i(t)}{R^i_{optical}} + 0 + \frac{L^j_{packet}}{R^j_{wireless}} + T^i_{j,prop-wireless} & \text{, if Case1a} \\ \frac{Q^i(t)}{R^i_{optical}} + (t^i_{next} - t) + \frac{L^j_{packet}}{R^j_{wireless}} + T^i_{j,prop-wireless} & \text{, if Case1b} \\ \frac{Q^i(t)}{R^i_{optical}} + (t^i_{next} - t) + \frac{L^j_{packet}}{R^j_{wireless}} + T^i_{j,prop-wireless} & \text{, if Case2,3} \end{cases} \tag{3.10}$$

In above equation, the first element calculates the queuing delay. The second element indicates the polling delay. The last two elements represent the transmission and propagation delay in the wireless transmission.

3.5.3 CAC in the Hybrid Optical Wireless Network

In order to support and protect the QoS of real-time traffic streams, in addition to bandwidth allocation, an AC scheme is required to decide whether to admit a real-time traffic stream based on both admission policies and QoS requirements supplied by the application at the end users. Assuming for real-time services, QoS metrics are predefined and various thresholds are specified. As mentioned earlier, the problem of the existed resource management in hybrid optical wireless networks is the lack of overall considerations to both optical and wireless network conditions. Admitting a new connection request in a WiMAX network, without considering the queuing time and polling time in the following EPON system will grant bandwidth to nonconforming traffic, e.g. to grant bandwidth to a service request, for which the delay requirements cannot be met.

Allocating transmission opportunities to nonconforming traffic will cause waste of scarce bandwidth and QoS degradations to conforming traffic. Intelligent decisions could be made to admit suitable traffic flows with guaranteed QoS in the network. Our proposed admission control algorithm is a rate-based Integrated Optical Wireless Admission Control (IOW-AC) scheme, by which the admission mismatch between the optical and wireless networks is avoided and the overall delay in the hybrid network is evaluated.

The real-time traffic flows are characterized by the demanded QoS parameters, for example, the delay bound (d_{min}). The AG can perform the admission control, which is based on the delay requirement of the new arriving flow and the estimated delay. The IOW-AC provides guaranteed QoS and increases the network throughput by accepting requests without violating the delay constraints. The proposed IOW-AC scheme can be accomplished as illustrated in Figure 3.5. When a new request received at AG, the transmission and propagation delay are first calculated according to the wireless network conditions and traffic profile. Secondly, the queuing delay is evaluated based on the current buffer occupancy. Then, using the polling status information, the expected polling delay is computed as explained in Equation 3.2 and Equation 3.6. Finally, when the IOW-AC approach determines the overall delay, d_{est}, for a multi-domain connection request. This value is compared to the delay with the delay bound, d_{min}. If the required delay bound is satisfied, the request is accepted ($d_{est} < d_{min}$). Otherwise, the request is rejected.

3.6 Performance Evaluation

Extensive simulations have been conducted to evaluate the performance of the proposed enhanced resource negotiation and integrated admission control for hybrid optical wireless network. A discrete event system, OPNET modeler [60] , is used to provide simulation results. Detailed simulation platform is given and results are presented to compare the proposed mechanism with the traditional admission control scheme. The simulations are carried out for different network sizes and different parameter settings.

The total number of AGs K is 32 and the EPON link rate is assumed as 1 Gb/s. The guard time between two adjacent transmissions on the optical uplink fiber is 5 μs. The upstream traffic from SSs to the OLT is considered and generated. The EPON MPCP operation is modeled. In WiMAX, the request/grant control mechanism is simulated. The traffic pattern is generated as Poisson distribution. Both real-time traffic

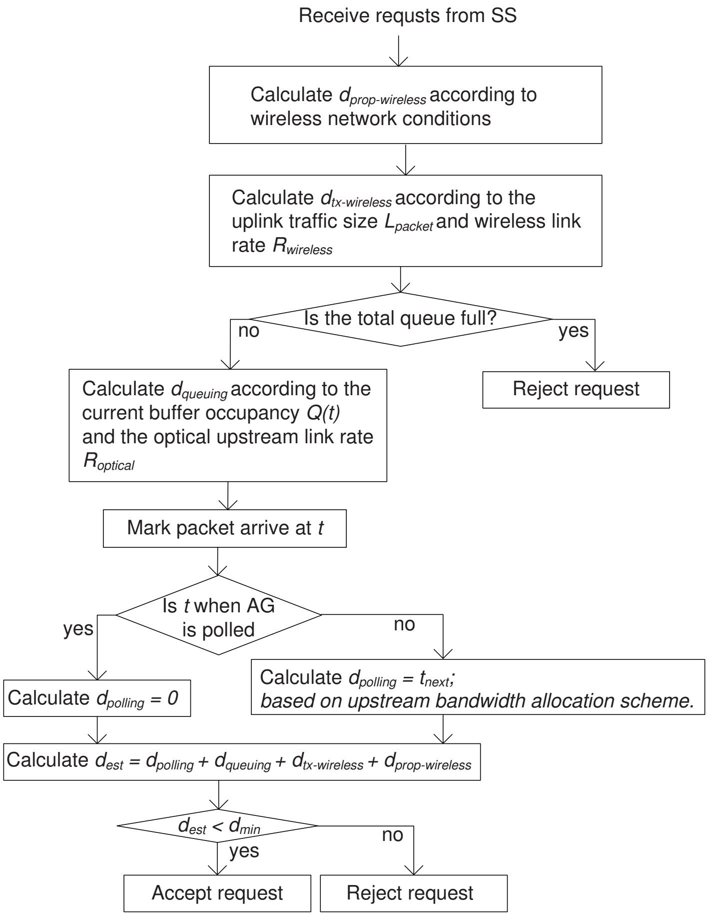

Figure 3.5: Flowchart of the proposed uplink IOW-AC scheme operations.

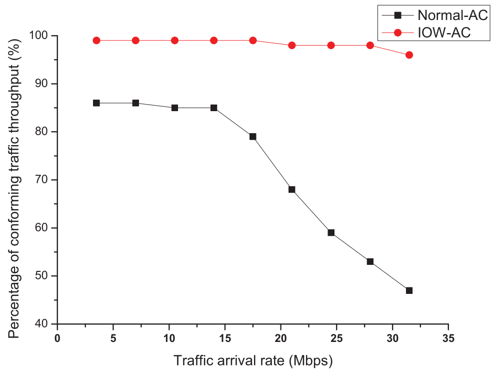

Figure 3.6: Comparison of IOW-AC and Normal-AC on the percentage of conforming traffic throughput.

and BE traffic are considered. For users of real-time traffic, e.g. video streaming service, the traffic profile has a variable packet size, uniformly distributed between 400 and 1500 bytes. The delay bound for the real time traffic is 75 ms. This threshold value is chosen as: in Reference [23] the end-to-end delay requirements for real-time application is less than 150 ms. In this work, the transmission covers from the user to the OLT, which is half of end-to-end transmission. Therefore, the delay threshold is 75 ms in this simulation.

Figure 3.6 shows the percentage of conforming traffic throughput for different traffic arrival rate and compares results between the IOW-AC scheme and the Normal-AC scheme. Results show that the IOW-AC shows better admission control ability to the nonconforming traffic. Being aware of the next cycle time, IOW-AC is able to take the polling delay in the EPON system into account and therefore block more nonconforming traffic than the normal AC. The IOW-AC shows superior behaviour especially when the real-time traffic arrival rate becomes high. The channel utilization becomes better in the IOW-AC scheme because there is wastage in the Normal-AC scheme taken by the nonconforming traffic.

Figure 3.7 shows the ratio of admitted requests in the total received requests for both real-time traffic and BE traffic, using the two different admission control schemes. The incoming requests that can be admitted only when their QoS requirements are met. This figure shows that the IOW-AC admitted larger amount of conforming traffic than the Normal-AC for both the real-time and BE traffic. The differences become more significant when the traffic arrival rate increases.

In Figure 3.8 the simulation results of the dropping probability are simulated for the real-time traffic and BE traffic, under a condition of limited buffer size at an AG node. IOW-AC scheme achieves lower dropping probability for both real time traffic and best effort (BE) traffic. This is because IOW-AC only accepts conforming traffic and saves more bandwidth for the conforming real time traffic without any drop even under high real-time traffic rate. It is observed that there are similar results as throughput performance. The IOW-AC scheme serves more BE traffic instead of dropping packets, which yields better bandwidth utilization.

In Figure 3.9 the percentage of accepted requests are plotted for different length of subframe period ranging from 1 ms to 5 ms. As been mentioned before, when the subframe period ($T_{granted}$) increases, there are more requests can be served in the uplink optical link. However, the total cycle time (T_{cycle}) becomes larger and the polling delay increases simultaneously. In this figure, the IOW-AC scheme blocks non-conforming traffic especially when the subframe period is large. Under the Normal-AC scheme, the accepted real-time traffic is most of non-conforming traffic type.

In Figure 3.10, the polling delay ($d_{polling}$) experienced at the AG node in the IOW-AC scheme is examined. The buffer size is specified as unlimited. Therefore, the $T_{queuing}$ of voice and video traffic will not be a dominating factor in the AC. On the contrary, the influence of $T_{polling}$ on AC is highlighted, which is important to observe in simulations. The value of the polling delay is zero if the request packet is received during the subframe period. Otherwise, the polling delay is the interval between the current time and the next poll time. In this figure, the polling delay is a time average value. As the subframe period increases, the polling delay becomes larger because it takes longer time for the OLT to serve all AGs. Therefore, the polling delay is also depending on the number of connected AGs. The more the number of AG, the longer the polling delay.

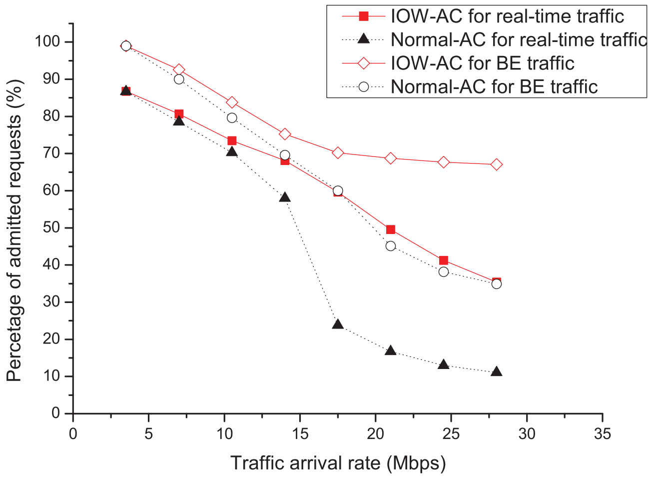

Figure 3.7: Comparison of IOW-AC and Normal-AC on the percentage of admitted request number.

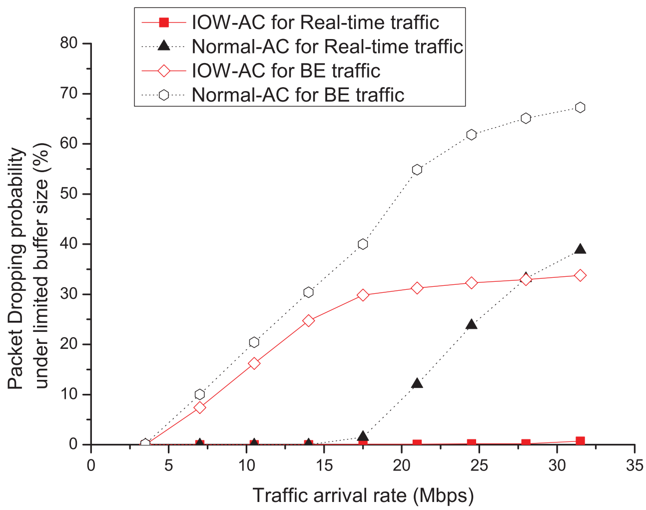

Figure 3.8: Comparison of IOW-AC and Normal-AC on packet dropping probability under limited buffer size.

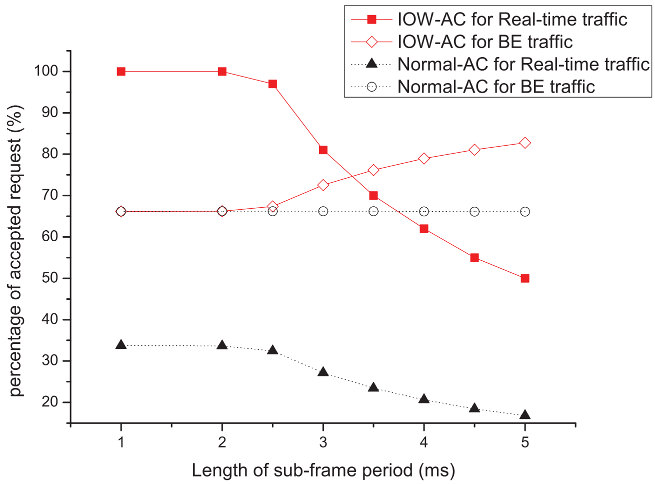

Figure 3.9: Comparison of IOW-AC and Normal-AC on accepted request under limited buffer size.

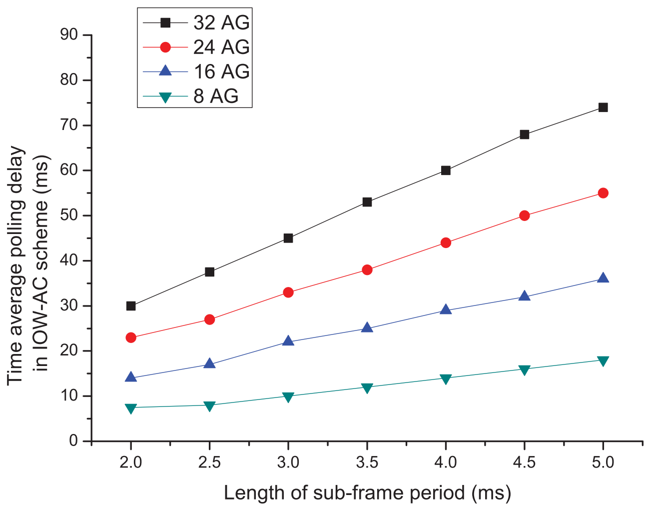

Figure 3.10: Time average polling delay in IOW-AC scheme under different number of AGs and different length of subframe period.

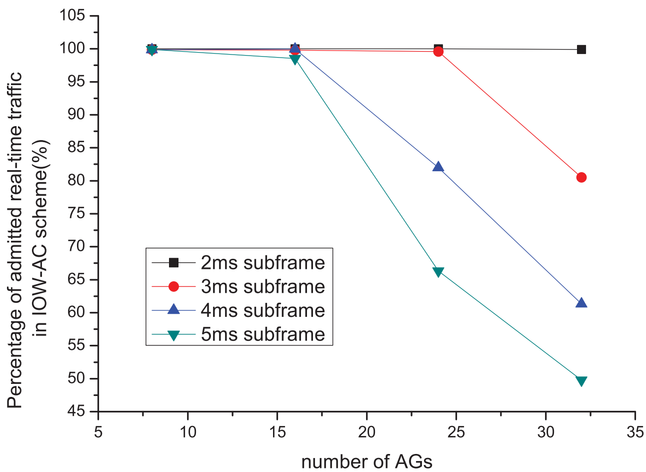

Figure 3.11: Percentage of admitted real-time traffic in IOW-AC scheme under different number of AGs and different length of subframe period.

In Figure 3.11 and Figure 3.12, the admitted real-time traffic and the BE traffic are compared under different number of AGs and different length of subframes, in the IOW-AC scheme. The number of blocked real-time requests increases using the IOW-AC scheme when the number of AGs and the period of subframe become larger. It is observed that the benefit of the proposed integrated resource management mechanism highly depends on the network size and traffic profile.

3.7 Summary

In this chapter, an integrated resource management mechanism is presented for the hybrid optical wireless networks aiming for maximum user QoS and maximum network throughput. The optical network and wireless network is jointly optimized using an integrated resource reservation and admission control scheme. The control platform aims to enhance the signaling communication with the least modification to the existed protocols in optical and wireless domains and the highest network channel

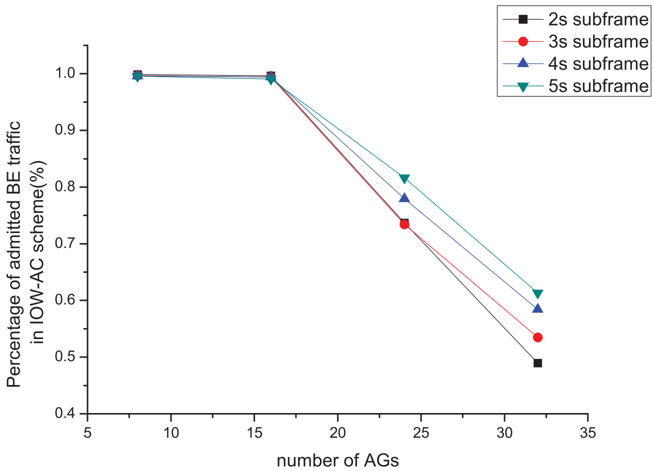

Figure 3.12: percentage of admitted BE traffic in IOW-AC scheme under different number of AGs and different length of subframe period.

utilization. The integrated resource management is favorable to provide QoS guarantees of real-time applications and improve the throughput of the best effort traffic. The main contributions of this chapter are as follows:

- An enhanced resource negotiation scheme, where the control messages in the existing EPON signalling protocol, MPCP, are extended with a field to carry the additional information of optical network conditions;
- An integrated optical wireless admission control algorithm, which makes decisions for the overall integrated network.

Simulations conducted using the OPNET modeler show that the proposed system achieves significant improvements over the traditional admission control approach in terms of the user QoS experiences and the network resource utilization. These gains are achieved without complex modification on the existed network protocols. Another interesting observation is that if the assigned subframe period for each AG are increased, the proposed admission control scheme achieves more significant improvements. Since the network throughput is relevant to the AG parameters, an operator can estimate the performance based on this simulation results. The benefit of the proposed integrated resource management mechanism highly depends on the network size and traffic profile. The analysis provides a good assistance or guidance in the research of hybrid optical wireless network architecture, including resource management, optimizing QoS and demanding service provisioning.

Chapter 4

Load Balancing Mechanism in the Integrated Control Platform

In this chapter, the problem of load balancing in terms of optimal cell selection and transmission power assignment is covered. The model applies for the convergence network of EPON and WiMAX where the TDM-like resource assignment scheme and the WiMAX air interfaces are used. Cell breathing is a well known cellular telephony concept that handles congestion by changing the coverage area of a fully loaded cell. In WiMAX, cell breathing can be achieved by changing the transmitting powers of the base stations. In the integrated EPON and WiMAX system, access gateways, i.e. EPON integrated base stations, report their residual downlink queue sizes to the central office using modified Multi-Point Control Protocol (MPCP). In this chapter, the cell breathing technology is studied as a load balancing mechanism into WiMAX network with an integrated EPON backbone. Centralized gateway selection and power allocation can be made jointly in the central office for all connected access gateways.

The major research issues are outlined and a cost function based optimization model is developed for power management. An iterative algorithm is proposed that equalizes the expected transmission time based on the reported queue sizes and adjusted transmission power level. In particular, two alternative feedback schemes are proposed to report wireless network status. For a given set of multimedia users with minimum quality of service requirements and a set of best effort users, the optimal resource allocation is decided, so that services for the users are guaranteed and the total network utilization is maximized. The proposed algorithm is simulated in integrated EPON and WiMAX system and the results show validation of the framework, i.e., load balancing schemes significantly outperforms fixed power control schemes.

4.1 Introduction

Optical networks offer greater bandwidth and reliability while wireless networks can support greater coverage and mobility. In hybrid network architectures, the resource management mechanism plays a key role to ensure an efficient usage of both optical link and radio spectrum. This is of particular importance for next generation broadband access networks which support simultaneously data, voice and multimedia services (triple play services) to multiple users. A system level control platform for integrated resource management is required to be considered since most of research efforts in this field have been focused in either optical networks or wireless networks solely. Our previous chapter shows demands for an integrated control plane to achieve overall and cooperated administration. Study of several system functions are of interest, including resource distribution, scheduling, and call admission control. This chapter continues the investigation of the integrated control platform in the hybrid EPON and WiMAX network. In particular, a load balancing scheme using the cell breathing technique is explored. The integrated network architecture is illustrated in Figure 4.1, which shows a mapping between clients and AGs. In the hybrid optical wireless network, the Optical Network Unit (ONU) functions and the Base Station (BS) functions are integrated into a single device, namely an *Access Gateway* (AG), which handles connections within the wireless network (single-domain connections), or cross both EPON and WiMAX (multi-domain connections). Clients associate with an AG with the strongest Received Signal Strength Indicator (RSSI) of the AG's beacon. Previous studies [61] [62] show that client service demands highly vary in terms of both time-of-day and location. Thus, traffic loads are often distributed unevenly among AGs, which results in congestions at popular locations.

As shown in the figure, AG1 can become overloaded while nearby AGs are lightly loaded.

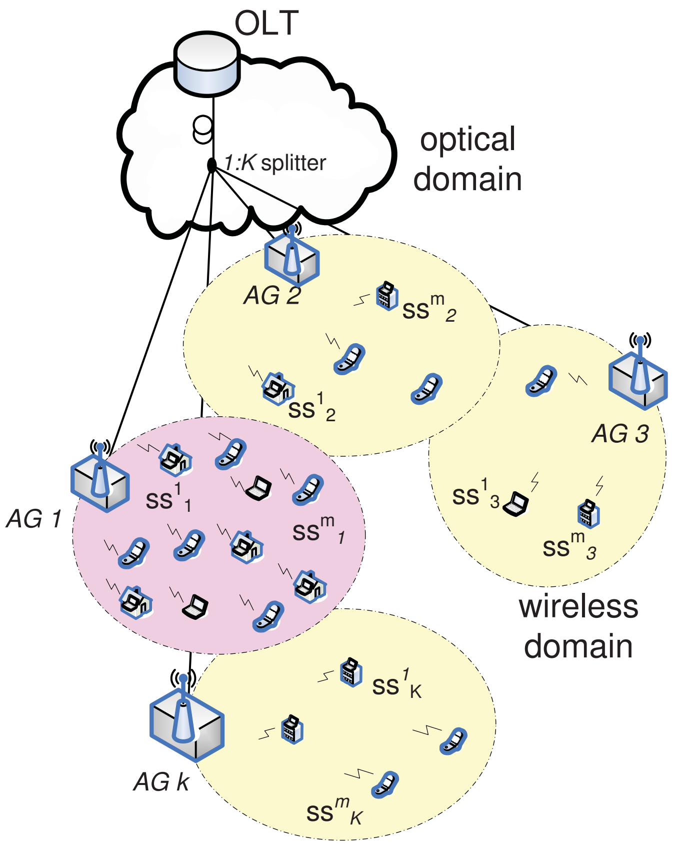

Figure 4.1: Illustration of imbalanced load distribution among AGs in the hybrid network. All AGs are assigned the same power value. AG1 becomes heavily loaded, as indicated by red color, while its neighboring cells have sufficient resources.

In an integrated optical and wireless network, coordinated control could be accomplished through the use of an integrated control framework. There are many opportunities to exploit the interworking between integrated EPON and WiMAX networks. The principles of the hybrid network architecture allow the OLT to distribute data to AGs, which further relay data to the target clients. To take advantage of the centralized system and the integrated control scheme, a load balancing algorithm and an optical downlink scheduler are proposed, which is able to control the AG transmit power dynamically and relocate traffic load in the edge area away from the heavily loaded cell. In load balancing, AGs are assigned with appropriate power level to avoid overload situations if a collocated system still has sufficient resources. A centralized power allocation control is implemented in the OLT unit. In each front-end wireless network, the base station is assigned with a proper power level to achieve quality transmission. This is done by exploiting a set of suitable power profiles that derive the maximum network throughput and user quality of service. The concept of cell breathing is used in conventional cellular networks to dynamically change the coverage area of cells. Fully loaded cells contract their coverage area whereas the lightly loaded cells expand their coverage area. Basic idea of load balancing is to relocate users in boundary regions from fully loaded cells to lightly loaded cells. Our goal is to propose a novel integrated load balancing scheme, which utilizes a cooperative signalling protocol to collect WiMAX network information and makes centralized power assignment decisions in the EPON. The remainder of this chapter is organized as follows. After presenting related work in Section 4.2, the system model is introduced and the optimization power allocation problem in Section 4.3. The relevance for MPCP-based load balancing mechanism is addressed in Section 4.4. This approach is validated by simulation results for an integrated network scenario is formulated in Section 4.5. Finally, this chapter is summarized in Section 4.6.

4.2 Relate Works

Several researchers have studied resource allocation in wireless networks in different environments, aiming at balancing load distribution and achieve efficient channel utilization. As suggested in previous works, one approach to addressing the load balance issue is to dynamically adjust the coverage area of a cell to reduce or increase loads. Because the finding of the appropriate power assignment to the cell tower to achieve load balancing is a challenging problem, proposed cell breathing algorithms typically rely on local heuristics [63–71]. Qiu and Mark [72] propose to use perceived Signal to Interference plus Noise Ratio (SINR) at cell to determine if the cell is sufficiently crowded. Du et al. [73] present a distributed bubble oscillation algorithm to achieve load balancing as a multidimensional resource allocation problem. Recently, there are also increasing efforts in applying cell breathing techniques to other wireless packet communication network such as WLAN [74] to achieve efficient resource utilization. While these schemes are able to achieve load balancing through the contraction/expansion of the cell sizes, they rely on heuristic algorithms because the cell breathing decision is made through distributed mechanisms. Besides adjusting the cell size, there are other ways to achieve cell breathing. Sang et al. [75] propose a cross-layer framework that enables call-level cell site selection and handoff. The algorithm is a coordinated one and utilizes a central server to make the scheduling parameter based on feedbacks on current cell loads. The loads are calculated based on the minimum rate requirement for each clients and the mean measured rate of each link. Instead of varying coverage area through the use of power control, the algorithm reduces the amount of time slots allocated to the users at the boundary.

Recently, there are significant interests in converging optical and wireless networks to exploit the complementary characteristics of the two networking technologies. These papers primarily focus on the architectures for integrated access networks and offer examples of potential enhancement. Shaw et al. [50] present a novel integrated routing method to achieve load balancing through the use of load-aware routing in multi-hop network. The approach is applicable to mesh network, but it does not provide the mechanism to achieve load balancing in single hop Point to Multiple Point (PMP) networks. To the best of author's knowledge, load balancing method with detailed protocol design has not been proposed in this area. This chapter takes a look at the challenging yet interesting multi-cell load balancing problem in view of network capacity maximization, by utilizing the integrated EPON backbone to make cooperative and centralized decisions for traffic distribution and power assignment. The proposed cell breathing mechanism is based on interworking between AG and EPON Optical Line Terminal (OLT) to dynamically adjust the number of Subscriber Stations (SS) associated to each AG.

4.3 System Model

We consider the hybrid architecture with multiple WiMAX cells as frond-end networks, where K AGs that connect to a central control station, OLT, via the optical link. There is one AG in each cell. As shown in Figure 4.1, the wireless network consists of a set of k cells $\{k = 1, 2, \ldots, K\}$ of and a set of m users $\{m = 1, 2, \ldots, M\}$ in each cell. The service area is partitioned so that each client connects to only one AG at any given time.

For simplicity of presentation, a system consisting of two cells is used to illustrate the cell breathing technique as shown in Figure 4.2. Cell operation area is modeled by a circle with a coverage radius (R_r). At any given time slot n, the cell coverage, the number of connected SSs, and the offered data rates to SS are varied by allocating different transmit power levels (P_t) to the AG. For instance, AG^1 initially is assigned with power P_t^1 and covers SS_{1-4}^1. After AG^1 is reported as overloaded, R_r^1 is contracted and its neighboring cell (R_r^2) is expanded in order to accommodate and serve additional SS_{3-4}^1. Thus, the traffic load is balanced between AG^1 and AG^2 with minimized packet loss.

Within a WiMAX cell all users share the common channel by using a Time Division Multiple Access (TDMA) radio interface. The delivery of the packets to subscriber users is performed on a TDM frame consisting of N slots. Mobility is not considered in our simulation scenario, only stationary (e.g. fixed terminals) or quasi stationary users (e.g. pedestrian) are assumed. The handover procedure is out of the scope of this work. We assume that fast handover is supported to minimize delay and delay jitter. In order to optimize the system and ensure overall QoS, the queue size and delay merit are used to determine whether or not a cell is overloaded. A cell overloading is said to occur, if 1) the total queued data for all traffic classes exceeds a threshold, Q_{thr}, or 2) delay to transmit the queued real-time traffic is violated to the requirement, D_{thr}. In the former case, the cell cannot take more traffic due to the buffer size limitation. In the latter case, the cell cannot guarantee real-time traffic with satisfied QoS requirements.

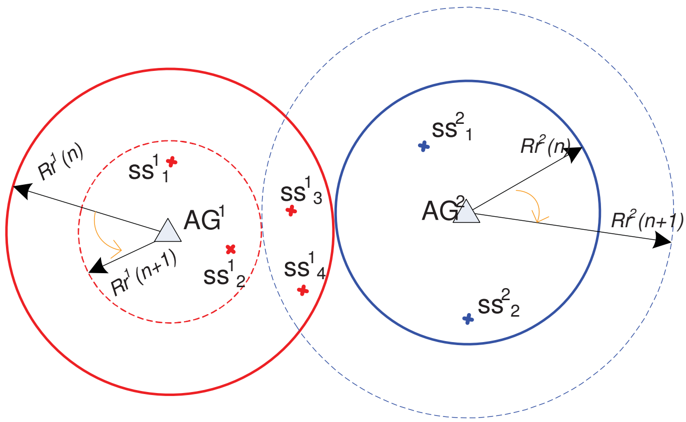

Figure 4.2: Cell breathing in wireless network. The solid circles represent initial coverage areas (at time n) and the dashed circles represent adjusted coverage areas (at time n+1).

4.3.1 Air-interface in Wireless Networks

The size of cell coverage depends on transmit power, noise and path loss. The received signal at the m^{th} user in the k^{th} AG, SS^k_m, is an attenuated version of the transmitted power level. The propagation model includes path loss, shadowing and fading. In this chapter, the propagation model defined in the IEEE 802.16 standard [76–78] is used to calculate the path loss (PL).

$$PL(dB) = 20\log_{10}\left(\frac{4\pi d_0}{\lambda}\right) + 10\gamma\log_{10}\left(\frac{d_0}{d}\right) + \Delta\mathrm{PL}_f + \Delta\mathrm{PL}_h + s \quad (4.1)$$

where λ is the wavelength. γ is the path-loss exponent. The parameter d is the distance between the BS and receiver. The parameter d_0 is 100 m. The parameter PL_f is the frequency correction term and the parameter PL_h is the receiver antenna height correction term. The parameter s is a log-normal shadow fading component.

The mean received power (P_r) is the difference between the sum of the transmit power (P_t) and the antenna gains (G_a) and the path loss (PL). The Signal to Interference plus Noise Ratio (SINR) is the ratio of received signal power to the interference plus noise signal power (I_n), which is computed by Equation 4.2. There is a minimum acceptable SINR ($SINR^{thr}$) for receiving a quality satisfied signal.

$$\mathbf{SINR}(dB) = P_r - I_n = P_t + G_a - PL - I_n \tag{4.2}$$

where I_n denotes the total interference received at each user, which consists of two parts: intra-cell interference and inter-cell interference. The intra-cell interference is caused by other SSs within the cell and the inter-cell interference is caused by the neighboring BSs. In this scenario, TDMA scheme is applied and only one SS transmits during the assigned slot time. It is assumed that there is no mutual interference among SSs in a cell. On the other hand, the initial cell coverages of AG^1 and AG^2 are non-overlapping by assigning proper transmission power values in order to avoid cross-cell interference. Our analysis is restricted to a static scenario, where it is assumed that all channel gains are constant. The theoretical system throughput of a Single Input Single Output (SISO) system, AG^k, can be derived from the Shannon capacity [79] using the expression in Equation 4.3:

$$T^k_{total} = \sum_{i=1}^{N} \log_2(1 + \mathrm{SINR}^i), \quad \forall k \tag{4.3}$$

Clearly, the throughput is determined by the interference level, the channel condition, applied transmitting power, and distance between the AG and SSs. For simplicity, it is assumed that the dynamic power control at AG allowing unique data rate for individual SS is not considered here. Thus, once the transmission power for the k^{th} AG during a slot time t, $P^k_t(t)$, is assigned, the system capacity can be estimated.

4.3.2 Service Queues

The main challenge for an integrated cell breathing algorithm is to choose the optimal AG association and combination of power assignments. At the same time, it is of importance to ensure transmission rate and delay. To formulate the queuing states in this integrated cell breathing problem, following notations are used:

$r^i(N)$: residual queue length in AG^i at the beginning of period N

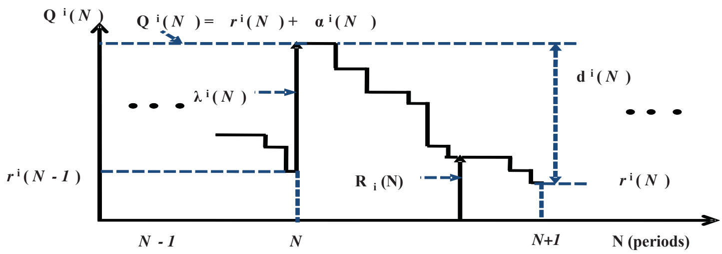

Figure 4.3: Downstream queue occupancy for AG^i during cell breathing control.

$\alpha^i(N)$: total traffic in AG^i arrived from OLT during period N

$d^i(N)$: total traffic departed from AG^i to the SS during period N

$Q^i(N)$: downlink queue size in AG^i at the beginning of period N

To visualize the problem, Figure 4.3 illustrates the AG^i queue occupancy under integrated cell breathing control. At the beginning of period N, a burst of packets $\alpha^i(N)$ arrive at the AG^i from the OLT and a control message may also be sent to the AG^i with the power assignment if power control is issued. At the beginning of period N, AG^i holds the amount $r^i(N-1)$ of data in the queue from previous period $N-1$. During this period N, the AG^i is expected to transmit $E[d^i(N)]$ to its SS. $E[d^i(N)]$ is denoted as $d^i(N)$ in the following discussion and the difference between $E[d^i(N)]$ and $d^i(N)$ are reconciled when the actual queue length, $r^i(N)$), is reported. The relationships between the queue size and residual queue length are:

$$Q^i(N) = r^i(N) + \alpha^i(N) \tag{4.4}$$

$$r^i(N+1) = Q^i(N) - d^i(N) \tag{4.5}$$

Thus, the difference equation for the AG^i queue length is:

$$Q^i(N+1) = r^i(N+1) + \alpha^i(N+1) = Q^i(N) - d^i(N) + \alpha^i(N+1) \tag{4.6}$$

In the proposed integrated cell breathing problem, the AG load is seen as the aggregate load contributed by its associated users. During a given network state, a subset of the AGs that suffer from maximal load is called the congested AGs. The objective of the load balancing algorithm is to minimize AG congestion load. The OLT determines the load of the AGs by allocating different power levels. The optimal controller would move the congested AGs into un-congested states and minimize the impact of the transfer loads to neighboring AGs. We refer to the resulting optimum queue length q^{i*} from the optimization, where the AG is neither under or over utilized under the optimal load allocation. To quantify the effectiveness of such cell breathing controller, a state variable $x^i(N)$ is introduced to quantify the difference between the actual queue length from the desired queue length, $x^i(N) = Q^i(N)$ - q^{i*}. The difference equation for $x^i(N)$ can be further derived as the following:

$$x^i(N+1) = Q^i(N+1) - q^{i*} = \{Q^i(N) - d^i(N) + \alpha^i(N+1)\} - q^{i*} \quad (4.7)$$

It is further defined $u^i(N) = \alpha^i(N+1)$ - $d^i(N)$, then there is:

$$x^i(N+1) = Q^i(N) + u^i(N) - q^{i*} \quad (4.8)$$

The variable $u^i(N)$ contains two quantities, $\alpha^i(N+1)$ and $d^i(N)$, to denote the traffic to be allocated to AG^i (control variable) and expected amount of traffic departed during current period N (measurement variable). The second parameter is not exact and in this formulation, the deviation between $E[d^i(N)]$ and the actual $d^i(N)$ is considered as a source of noise. The expected amount of departed data can be estimated using theoretical system throughput.

4.3.3 Problem Formulation

After the definition of the air-interfaces and service queues including the constraint, the optimization problem is now formulated. The cost that an AG suffers from assigned resources can be measured. System cost merits contain the system throughput, delay, and the backlogged data size. From the operator's perspective, the profit increment represents minimizing the total cost over all possible resource assignments. The objective is to assign resources so that:

- The sum throughput of all AGs is maximized;

- All connections with assigned transmission power deliver quality services;
- The congested load in AGs is minimized.

The output of the optimization is a table consisting of the cell that the downstream data should transmit to, and the power that is assigned for the corresponding AGs. The cost is a measure of how efficiently resources can be exploited if an AG is assigned with a certain amount of downstream data packets. Given a finite horizon of n frame, an optimal cell breathing controller would maximize the total network throughput, defined as Equation 4.9. Each AG^k attains its individual fixed transmitting power level during a time slot, P_t^k, to provide services with satisfied QoS constraints of each user, while optimizing the global multi-cell system throughput through balancing traffic load between AGs. Define $p = [p_1, p_2, \ldots p_{max}]$ as the vector of cell power levels assigned to AGs by the OLT.

$$maximize \quad \sum_{i=1}^{K} T_i(p_i) \tag{4.9}$$

Subject to

$$p_1 \le p_i \le p_{max}, \quad i \in K, \tag{4.10}$$

$$\text{SINR}_{ij} \ge \text{SINR}_{ij}^{thr}, \quad i \in K, j \in M \tag{4.11}$$

$$Q_i \le Q_{thr}, \quad i \in K, \tag{4.12}$$

In the above program, the first constraint indicates that each AG is assigned within a maximum power p_{max}. The second constraint shows that quality transmission is ensured with assigned power. The last constraint represents that the expected traffic load is less than a pre-determined congestion level (e.g. queue size limitation Q_{thr}).

4.3.4 General Iterative Algorithm

When a data frame arrives at the OLT for an SS, the central scheduler has two decisions to make: 1) find the associated AG to transmit the packet and 2) choose the minimum transmitting power to setup a qualified transmission from the associated AG to the destination SS. It is necessary for AGs to report information of their connected SSs to the central control office, the OLT. Required information includes the wireless channel condition (e.g. link rate), the distance between a user to the base station, the residual buffer size (Q_{res}), and the Residual Expected Wireless Transmission Time (REWTT). Upon receiving feedback from an AG, the OLT examines the load volume, evaluates congestion level, and assigns a value of transmit power to the AG. Having defined the cost minimization problem and derived the power assignment rules, an iterative algorithm is presented, in order to solve the optimization problem in centralized OLT unit. Each iteration consists of four steps:

1. If the cell is lightly loaded, such as $Q^i_{res} < Q^i_{thr}$ or $REWTT^i < D^i_{thr}$, the power levels remain their latest value without any change. Otherwise go to Step 2

2. When Q^i_{res} or $REWTT^i$ exceeds the threshold values, the cell is sufficiently crowded and the step 1) is executed for neighboring cells. The procedure finds available nearby cells, which have enough bandwidth to accept additional traffic load. If there is no such cell, then go to Step 4. Otherwise go to Step 3.

3. The transmission power of the highly loaded cell is reduced, and at the same time, the power levels in the lightly loaded neighboring cells are increased. Some users in the boundary region will be forced to handover from the overloaded cell to neighboring cells. There is a minimum power level, so that the transmission power cannot decrease unlimited.

4. If all neighboring cells unfortunately cannot assist, the power values will not be changed. Packet dropping may occur in the congested cell.

In Step 3, the OLT needs to decide the size for the highly loaded cell to shrink its range and for neighbouring cells to expand. In addition, a set of users in the boundary region is chosen to change their serving AGs by handover. The rules to select SSs are discussed in next section.

4.4 Proposed Load Balancing Mechanism

As previously introduced in 3.4, EPON relies on Multi-Point Control Protocol (MPCP) that is based on *GATE* and *REPORT* messages to grant and request for uplink bandwidth. The GATE message is broadcasted to all ONUs and target ONU receives the packets based on the labeled Link Layer Identification (LLID). The REPORT message is transferred to the OLT by ONU within its uplink transmission window. OLT allocates uplink bandwidth based on either fixed bandwidth allocation or Dynamic Bandwidth Allocation (DBA) algorithms. Here in the hybrid architecture, ONU functions are implemented within the AG unit.

The load balancing mechanism is implemented based on the traditional MPCP framework. In order to achieve load balancing, the cell breathing mechanism extends the MPCP control and incorporates two frames to achieve power control and load feedback. The two frames are *GATE-p* and *REPORT-p*. Similar to GATE, GATE-p is broadcast downlink and contains the LLID and P_t to indicate allocated power level to individual AGs. The GATE-p message is used for both assigning power information and upstream bandwidth allocation information. REPORT-p is a modification of the existing REPORT message and adds the residual queue length (Q_{res}) and Residual Expected Wireless Transmission Time ($REWTT$) to the message.

4.4.1 Information Exchanged and Cell Selection

Since the OLT has global knowledge of the state of the entire network, the process of power assignment and traffic distribution is globally optimized. Monitoring wireless network status is a mandatory feature in the load balancing mechanism in order to guarantee correct and sufficient knowledge at the OLT. Each AG is required to report information of connected SSs, including position (r_k), downstream bandwidth request (b_k, which is represented by the backlog information), and transmission data rate (R_k) of each SS.

Within a highly loaded cell, two factors are taken into account when choosing SSs to balance traffic load. First, SSs located at the edge of a highly loaded cell are considered. The OLT starts to scan SSs with low SINR (to its corresponding AG). If multiple users have the same distance, all of them are chosen because they are forced to be adjacent cells at the same time, once the cell range is reduced. All SSs, **S**, are arranged in an order based on their distances to the center base station in the highly loaded cell. A set of h SSs are chosen to be disconnected and reallocated:

$$S^* = [SS_1, SS_2, \ldots, SS_h], \quad where\, r_1 \geq r_2 \ldots \geq r_h\, and\, S^* \in \mathbf{S} \tag{4.13}$$

Second, the number of SSs which are redirected to other cells should be taken into account due to the handover overhead. The more SSs the longer time is required to reconnect to other cells. The total bandwidth request for S^* to be sent to other cells should be as close as possible to the difference between Q_{res} and Q_{thr} (or difference in mean delay), because impressing large bandwidth demand to the neighboring cells might cause congestions in the cells in the future.

$$\sum_{i=1}^{h} b_i = Q_{res} - Q_{thr}, \quad i \in h \tag{4.14}$$

$$\sum_{i=1}^{h} \frac{b_i}{R_i} = REWTT_i - D_{thr}, \quad i \in h \tag{4.15}$$

The implementation presented in next section considers a set of discrete and restricted power levels in order to reduce computational complexity and maintain sufficient interference control, respectively. After the AG selection and power assignment are completed, the OLT labels the messages with the appropriate target address and conveys to each AG via MPCP control message.

As presented earlier, congestion in a cell, if observed, is solved by contracting cell coverage and reallocating SSs to neighboring cells. In order to determine the size of reduced cell range, the OLT requires detailed information of SSs from each AG. To take the control message overhead into account, it is impractical to report the current queue status for each user. In order to scale the amount of control messages, a partitioning of the wireless coverage area is considered by dividing this into several inner circles. As shown in Figure 4.4, the total cell range is divided into a set of regions, $A = A_1, A_2, \ldots A_\alpha$, where α is the fraction number to fragment the maximum cell radius. Within an inner circle,

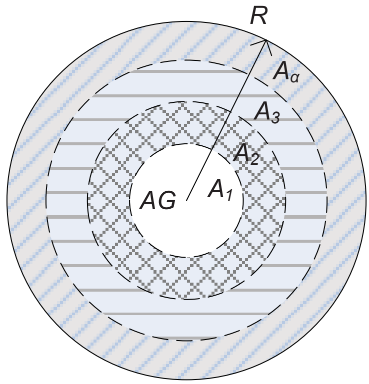

Figure 4.4: Partition of the wireless coverage area into inner circles.

the wireless link rates (considering the wireless channel condition) are averaged, and the queued data size is aggregated. Thus, the information embedded into the REPORT-p message is classified and simplified. The granularity of cell fragment decides the number of residual queue size and expected transmission time embedded in the REPORT-p message. The transmission power cannot decrease indefinitely. There is a minimum value and the power level is decreased step-wisely according to the division of inner circles. It is worth noticing that there is a trade-off between the number of reported queues and the additional overhead of REPORT-p message.

4.4.2 Feedback Schemes for the Load Balancing Mechanism

In order to collect network status from all connected AGs, two cell information feedback disciplines are developed and their performances in the MPCP based load balancing mechanism are compared. The first discipline is called Polling Period Report (PPR) scheme, where cells update their network status whenever they are polled by the OLT. We show that without increment of complexity compared to traditional MPCP, the scheme attains improved performances in terms of network throughput and delay. The second feedback discipline is called Short Period Report (SPR) scheme, which is an extension to the PPR scheme. We describe these two feedback schemes in the following subsections and then illustrate their performances by simulation results.

Polling Period Report (PPR) feedback scheme

Using the GATE-p message, granted bandwidth and start time are assigned to AGs. AGs are polled in sequence based on the scheduling policy used in the OLT. In this example, AGs are polled in an increased order of their LLID number using Round Robin (RR) scheduling. The following load balancing procedure is carried out (illustrated in Figure 4.5):

- At time t_1, the OLT broadcasts power assignments to its connected AGs via GATE-p message. Each AG adjusts its transmitting power and associates SSs.

- After receiving the GATE-p message, AG^1 is polled and starts its uplink transmission at t_2. Along with data, the current Q_{res} and $REWTT$ information of the AG^1 are reported to the OLT. Upon receiving the REPORT-p message at t_3, the initial entry table is updated at the OLT.

- If the AG^1 is highly loaded, the OLT reduces the assigned power to AG^1 ($Pt_2^1 < Pt_1^1$) and increases its neighboring cell power ($Pt_2^k > Pt_1^k$). Traffic load is examined for neighboring cells. If one neighboring cell is also highly or near highly loaded, the power value cannot be increased. In this case, some SSs are not able to handover to a new service BS.

- The GATE-p message is broadcasted along with the GATE-p message at t_4, and all AGs adjust their transmission power levels. Since the subscriber device chooses the base station with the strongest Received Signal Strength Indication (RSSI) among all received signals from AGs. When the cell coverage changes, some subscribers in the boundary region (e.g. SS_3^1, SS_4^1 in Figure 4.1) will be forced to handover to a less loaded neighboring cell.

- The GATE-p message destined to AG^2 contains both information for transmission power adjustment and information for upstream

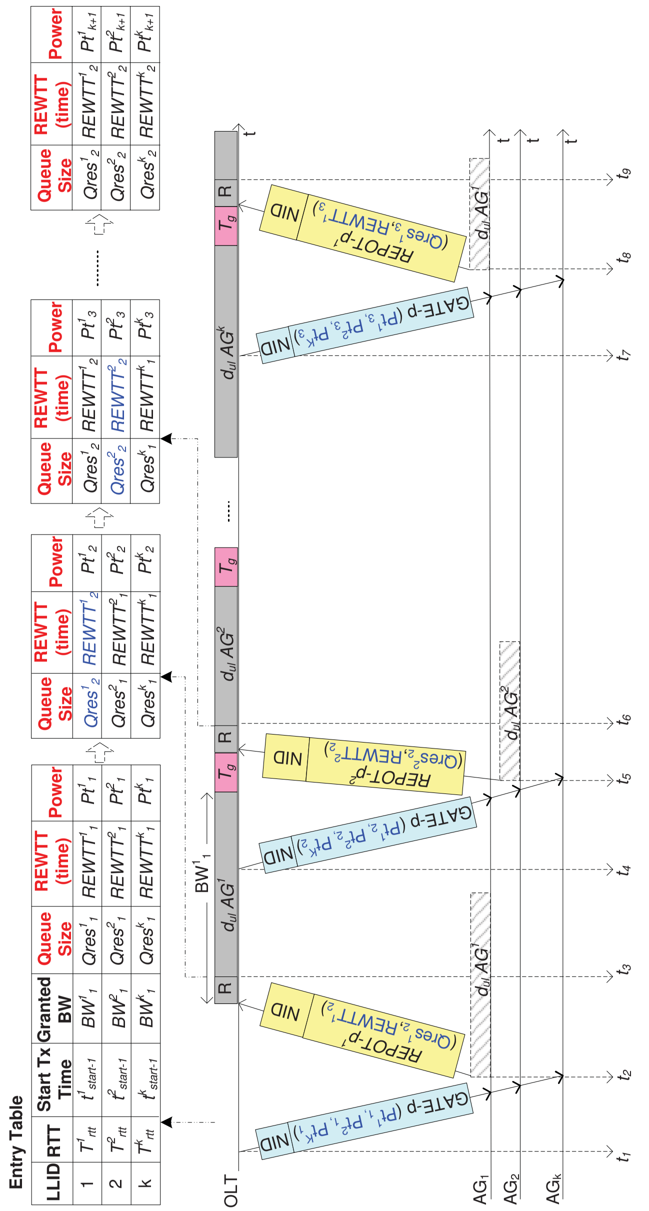

Figure 4.5: Load balancing mechanism with the Polling Period Report (PPR) feed-back scheme.

bandwidth allocation. At t_5, only AG^2 is granted with upstream transmission bandwidth.

- The new network status of AG^1 is updated again when AG^1 is polled again by the OLT. If there is power adjustment during the polling cycle (from t_2 to t_8), the feedback from AG^1 after decreasing/increasing its transmission power and contracting/extending its region is reported to the OLT at its next polling time, t_8.

As explained above, in the PPR scheme, an AG reports its current traffic load and channel conditions to the OLT only at the time the AG is polled. For example, the AG^1 changes its power at time t_2, and the consequence of power adjustment is learnt at the OLT at time t_9. When the fixed upstream bandwidth allocation (i.e. TDM) is used, the polling interval for an AG is fixed, so that the interval for each AG to report is also a fixed amount of time. When DBA is employed instead of fixed bandwidth allocation, REPORT-p message and the embedded downlink load information is fed back to the OLT aperiodically. As discussed in Section 3.4, a polling cycle (T_{cycle}) refers to a period in which all AGs are served. Along with an increasing number of AGs, waiting for a polling cycle to update the cell status may either result in a congestion that cannot be discovered on time, or, alternatively, yield improper power adjustment due to out-of-date information. Particularly, to account for random wireless channel fluctuations, the feedback mechanism is important to attain real-time cell status.

Short Period Report (SPR) feedback scheme

Although the implementation of PPR scheme is simple, the load balancing may be not optimized and effective because precise network status cannot be attained, especially when the polling interval for an AG is increased. Therefore, a feedback scheme with short updating period is proposed. This scheme can be implemented by coordinating the upstream bandwidth allocation at the OLT and AGs.

The basic idea of SPR scheme is to grant AGs upstream bandwidth to report their cell information after an AG is polled. As indicated in Figure 4.6, the OLT broadcasts the GATE-p control message to all AGs and assigns transmission power value according to the cell breathing algorithm to attain global network throughput optimization (at t_1). Upon receiving the GATE-p message, all AGs are required to inform their network conditions by transmitting the REPORT-p message (during t_2 to t_3). The OLT allocates a period for receiving REPORT-p messages from all AGs and an upstream transmission period for the polled AG. Different from the PPR scheme, the entry table maintained in the OLT is updated every granted slot time (at t_4 and t_8) to each AG instead of being updated after all AGs are granted.

Compared to PPR scheme, the period in the SPR scheme that the OLT receives network updating information is decided by the granted upstream transmission bandwidth to each AG using SPR scheme. Cell status is updated more frequently, so that the power management decisions are made according to nearly real-time network conditions. The period that the OLT allocates for receiving REPORT-p messages from all AGs is defined as the REPORT-p window. It is worth noting that the length of the REPORT-p window is determined by the number of connected AGs and the size of REPORT-p message. Obviously, the upstream link capacity is reduced by introducing the overhead of REPORT-p window. Next section will show the increment on network performance as a trade-off of introducing more control message overhead in the SPR feedback scheme.

It is well known that the uplink optical bandwidth is shared by multiple AGs in the hybrid network architecture. The most important operation at the OLT is to arbitrate the time for sending REPORT-p messages by allocating upstream transmission periods precisely to AGs. Using the PPR scheme, it requires no change on the uplink bandwidth scheduling function at the OLT. On the other hand, the SPR scheme requires two modifications:

1. The length of REPORT-p window (L_{RPWin}) is the sum of the transmission time of REPORT-p from every AG (S_{RP}). For the polling cycle of an AG ($T_{cycle-SPR}$), the calculation takes account of this newly introduced REPORT-p window.

$$L_{RPWin} = \frac{\sum_{i=1}^{K} S_{RP}^{i}}{R_o} \tag{4.16}$$

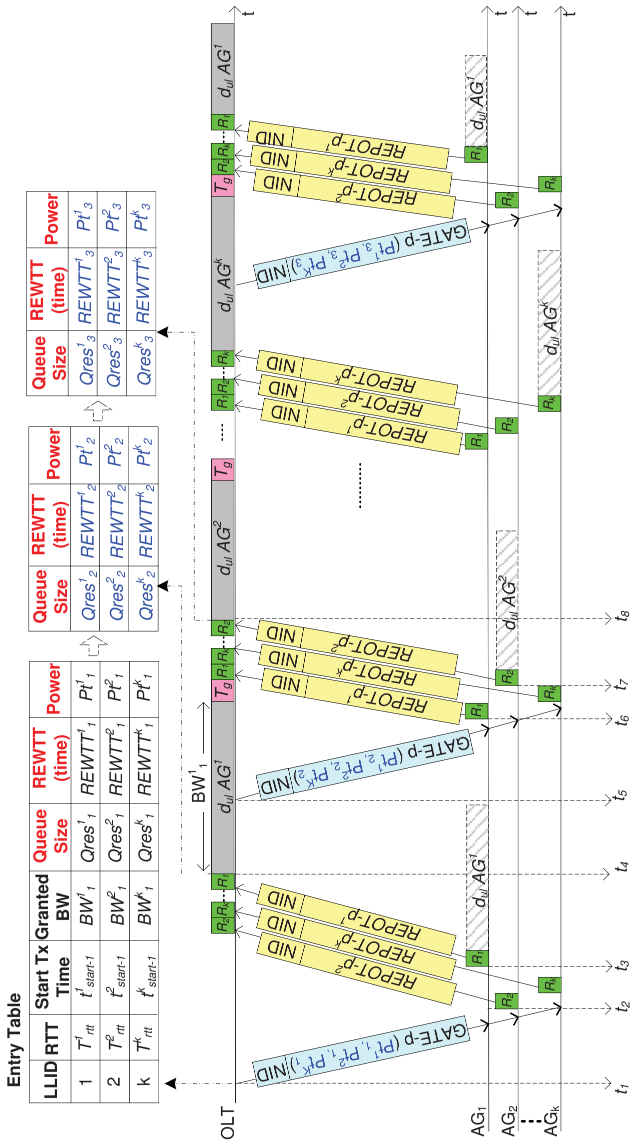

Figure 4.6: Load balancing mechanism with the Short Period Report (SPR) feedback scheme.

$$L_{cycle-SPR} = \sum_{i=1}^{K} \left(\frac{BW^i}{R_o} + T_g + L_{RPWin} \right) \quad (4.17)$$

where BW^i is the granted bandwidth to the AG^i and T_g is the guard time.

2. For the reason that the grant start time is used for transmitting both the REPORT-p message and data payload in the polled AG, the polled AG starts its transmission after other AGs, so that REPORT-p messages from other AGs are not delayed because of data payload transmission. The start time of the polled AG is calculated by Equation 4.18.

$$t^i_{start} = t^i_{grant} + \frac{\max\{T^1_{rtt}, T^2_{rtt}, \ldots, T^K_{rtt}\}}{2} + \left(L_{RPWin} - \frac{S^i_{RP}}{R_o} \right) \quad (4.18)$$

where t^i_{grant} denotes the time to transmit the GATE-p message. The second term on the right-hand side of the above equation selects the largest return trip time value, because the GATE-p messages are delivered to all AGs. The third term indicates that the polled AG is the last one to start upstream transmission.

4.5 Simulation Results

4.5.1 Simulation Scenario

In this section, the integrated load balancing using cell breathing scheme is evaluated in OPNET simulation environment [60]. The integrated system is similar to Figure 4.1 and consists of K AGs (K = 16, 32, 64, or 128). The downlink optical transmission rate is 1 Gbps and downstream data are broadcasted from the OLT to AGs. There are up to 200 SSs connected into a wireless network. Traffic arrival process follows a Poisson distribution and the mean arrival rate varies from 0.01 to 0.1 Mbps per SS. The guard time is 5 μs and the AG buffering queue size is 50 MB. In the WiMAX system, network parameters are configured

Parameters	**Value**
System	TDMA
Frequency band	3.5 GHz
System bandwidth	10 MHz
Path loss model	Erceg model [78];
Propagation environment	typeC: flat terrain with light tree densities [78];
User distribution	Uniform distribution
BS transmit power	40 dB
Fading st. dev	8 dB
BS antenna gain	17 dB
SS antenna gain	0 dB
BS antenna height	35 m above ground
SS antenna height	2 m above ground
Noise power	−138.41 dB
Noise figure	7 dB

Table 4.1: A summary of WiMAX system parameters

as in Table 4.1 [80, 81]. For simplicity there is one AG per wireless cell and SSs are distributed randomly over the cell region.

For the simulation study, one heavily loaded cell at AG^1 is considered. The neighboring cells of AG^1 are lightly loaded. Thus, the neighbor cells of AG^1 are available to load balance bandwidth from the boundary region of the AG^1. The simulation compares the performances of the load balancing (LB) solution against a scenario without using load balancing (NLB). In the simulation, AG drops incoming packet when either the queue is full or the delay requirement for real-time traffic is not met.

4.5.2 Test Case1: Comparison of LB and NLB mechanisms

We begin by testing the behavior of our load balancing scheme by varying the network traffic load. The scenario consists of 16 AGs. The number of SSs is fixed as 200 SSs per AG. There are up to 80 percent SSs located in the area, where is possible taken over by neighboring cells. The admitted downlink rate is varied, so that the overall traffic load increases incrementally as higher rate is assigned. Under different amount of allocated bandwidth, various performance merits including network throughput, average delay, and dropping packet size are measured.

We first simulate the network throughput rate vs. input traffic amount (shown in Figure 4.7). Input traffic amount is used instead of the input optical rate because optical line rate is significantly higher than wireless data rate. Thus, total input traffic amount is presented and the simulation generates the indicated amount of input traffic periodically over the entire observation period. As expected, when the load balancing scheme is applied, overloaded traffic in the AG^1 can be shared by neighbouring cells, and therefore, the overall network throughput increases. The LB solution increases the network throughput in excess of 25% at best and the improvements become more obvious when the input traffic load increases. It demonstrates that the proposed cell breathing operation can effectively handle heavy load in AG^1 by reducing the transmit power of AG^1 and reallocating boundary SS. Hence, with the load balancing of network, the congestion of the cell and in turn the whole system is effectively relieved. However, the total output throughput in the heavily loaded cell is decreased due to reduced transmission power. When input traffic load becomes heavier, AG^1 cannot shrink the cell range since there is a minimum power bound. In the case of NLB, AG^1 continues to serve the boundary SS, which overloads the AG while demanding relatively longer transmission time due to low SINR available at the crowded cell boundary. Hence it is advantageous to balance traffic load among cells.

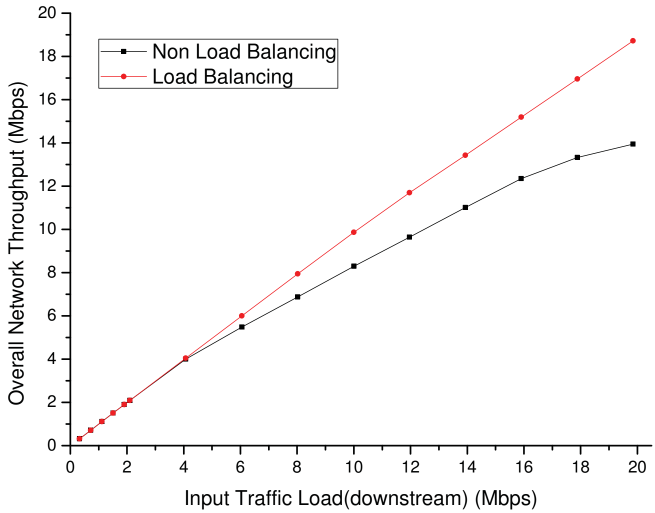

Figure 4.7: Comparison of LB and NLB mechanisms - overall network throughput

Next shown in Figure 4.8, the amount of dropped packets at AG^1 is evaluated under different input traffic load. Under light traffic load, the dropping probability of LB solution is zero, which performs better than the NLB solution. However, when the input traffic load continues to increase, neither LB nor NLB can meet the delay requirement or the maximum buffer size. The gap between the two curves consists of the output throughput rate difference and the traffic load from the boundary area. We note that LB algorithm outperforms NLB and the amount of improvement depends on the availability of boundary SSs that can be reallocated to neighboring cells.

Then the average transmission delay in the AG queue is investigated. The end-to-end transmission delay consists of the queuing delay in the

OLT, the transmission delay in the optical downstream link, the queuing delay in the AG, and the transmission delay in the wireless downstream. Because sufficient optical resources are assumed and high speed optical link is available, the queuing delay in OLT and the optical downstream transmission delay are neglected. The queuing delay in the AG is mainly determined by two factors: the queue length and the scheduling scheme. Using load balancing algorithm, the queue length in AG is affected, which is our focus. The benefit of using advanced scheduling schemes is discussed later. In this test case, the First-In First-Out (FIFO) scheduler is deployed for downstream transmission in wireless networks. Shown in Figure 4.9, when the input traffic load is moderate, using the load balancing algorithm achieves reduced queuing delay. This is because AG^1 is relieved from excess loads after load balancing and can more efficiently handle the remaining queued data. Therefore, using load balancing improves the delay performance.

4.5.3 Test Case2: impact of scheduling policies

We now show the impact of different scheduling policies at AG for downstream transmission in wireless networks. We choose a priority based scheduling policy, High Rate First (HRF), as a comparison. Simulation results show these measurements with LB and without LB (using FIFO and HRF schedulers). The HRF scheduler improves network performances by favoring connections with higher data rates when congestions occur in wireless networks. Note that by using load balancing scheme, congestions are avoided. Therefore, results in terms of network throughput and dropped packet size are similar under FIFO case and HRF case when LB is utilized (Figure 4.10 and Figure 4.11).

As for average delay performance in Figure 4.12, the HRF scheme outperforms the FIFO scheme in both NLB and LB cases. On the other hand, when the HRF scheduler is applied, the load balance scheme does not reduce transmission delay. This is due to the fact that when more input traffic load are admitted and AG power value is decreased, the bandwidth that was assigned for the high data rate connections is now shared by connections with lower data rates.

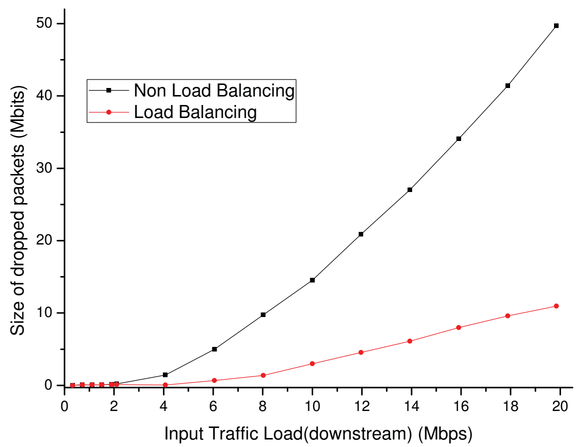

Figure 4.8: Comparison of LB and NLB mechanisms - size of dropped packets

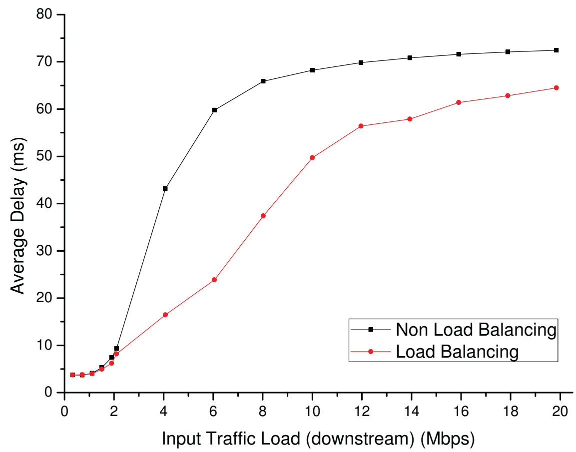

Figure 4.9: Comparison of LB and NLB mechanisms - average delay

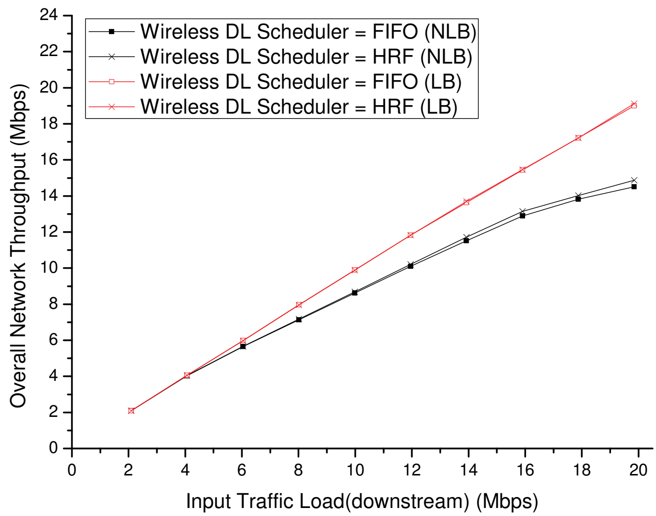

Figure 4.10: Impact of scheduling policies - overall network throughput

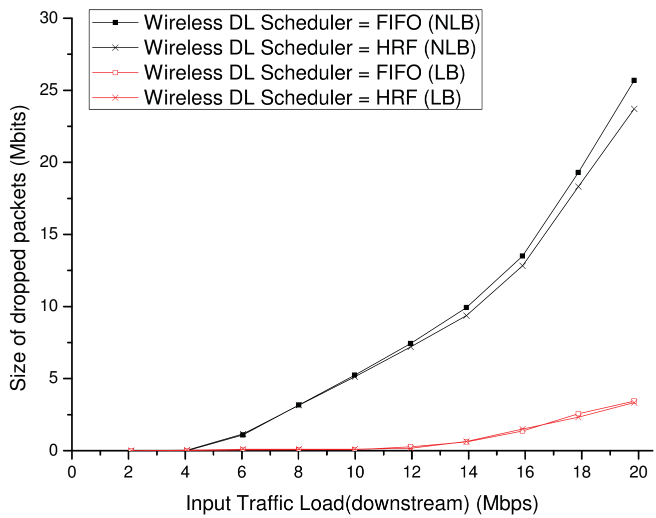

Figure 4.11: Impact of scheduling policies - size of dropped packets

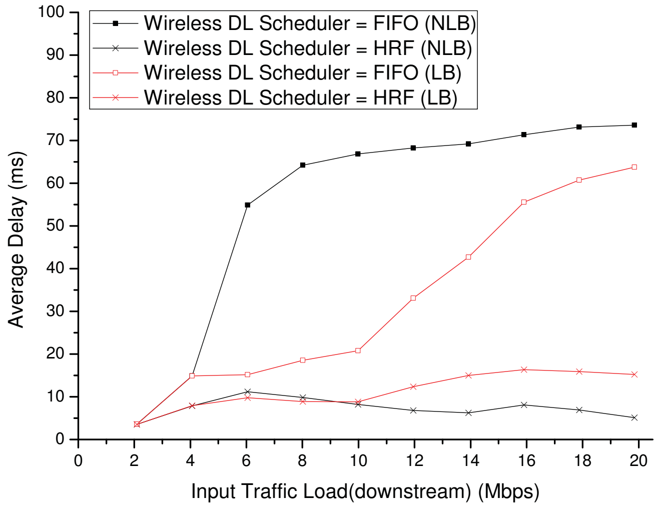

Figure 4.12: Impact of scheduling policies - average delay

4.5.4 Test Case3: impact of multiple feedback queues

We further investigate our load balancing scheme by measuring network performances with various numbers of queues, which are reported to the OLT via REPORT control messages. Although a small number of reported queues generates low overhead of control messages, it will result in an inaccurate estimation of the power adjustment for a cell. For example, the AG power changes in fine granularity and more queue status are required to be reported. If the AG is overloaded and has more traffic to transmit, then the OLT allocates a reduced transmission power and remove a certain traffic load. Unlike larger number of reported queues, where for a given overloaded cell the OLT determines total connections that should be shifted and the coverage range that should be reduced, for the case where insufficient queue information is provided, if more coverage range is decreased than necessary, then more users are forced to be redirected, and this results in wastage of the bandwidth utilization in the cell and high risk of causing congestion in neighboring cells.

We run simulations for aggregating queue status into 2, 4, 8, 16, and 32 groups and reporting to the OLT. The number of redirected users is shown in Figure 4.13. When report group is 2, half of SSs in the boundary area are removed even if only 10 percent of them are necessary. As for a report group with large number, it makes sure to satisfy the guaranteed bandwidth by removing suitable number of SSs. When the traffic load is increased, the number of redirected SSs becomes higher. However, when the system reaches saturation, the number of redirected SSs starts decreasing. This is due to the fact that when more incoming traffic is admitted, the cell becomes overloaded and contracts its coverage sooner. Accordingly, as illustrated in Figure 4.14, results show that there is at most 40% increment in network throughput. This, in real and practical settings, will cause a trade-off between the overhead of REPORT control messages and network transmission efficiency.

4.5.5 Test Case4: impact of multiple AGs

In this scenario, performance of the load balancing scheme with different number of AGs is analyzed. When the number of AGs is increased, the interval between two successive polling control messages for an AG is increased. In other words, the period of updating the cell status (i.e., queue sizes and channel conditions) is increased, which may cause problems if the out-of-date information is utilized during load balancing. Figure 4.15 depicts the amount of redirected packet size from the heavily loaded cell to neighboring cells. We observe that the volume of traffic is less when the OLT can derive information of the heavily loaded cell more frequently. The difference in the power adjustment and the number of reallocated SSs is due to the difference in the frequency of updating cell status. Subsequently, performances in terms of throughput, dropped packet size, and delay are simulated. From Figure 4.16 to Figure 4.17 it is observed that the results for the 16 AGs case are better than higher AG numbers. The load balancing scheme is more effective when the updating period is moderate.

4.5.6 Test Case5: Comparison of PPR and SPR feedback schemes

In this test case, performances of PPR and SPR feedback schemes are shown by simulation results. As discussed in Section 4.5.5, the network performances are degraded when the number of AGs is increased. The

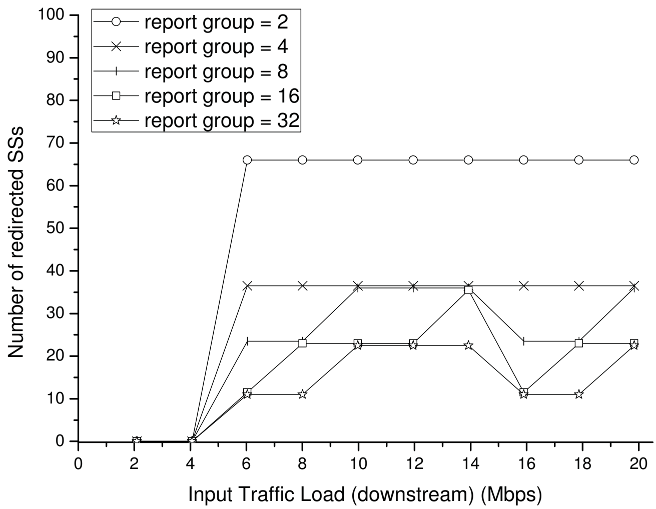

Figure 4.13: Impact of multiple feedback queues - size of redirected SSs

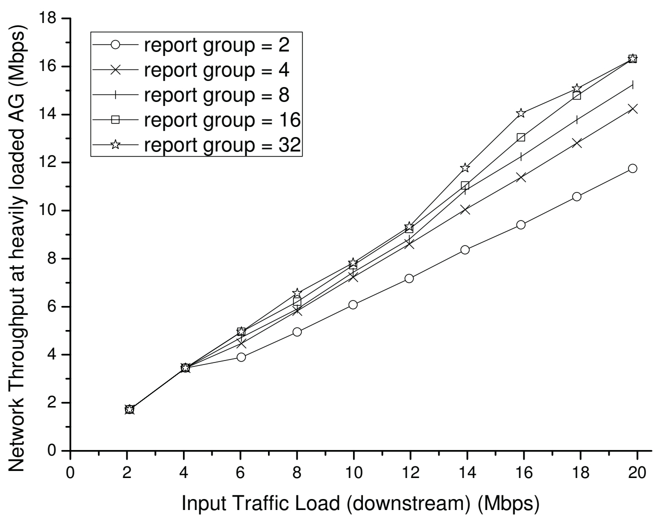

Figure 4.14: Impact of multiple feedback queues - overall network throughput

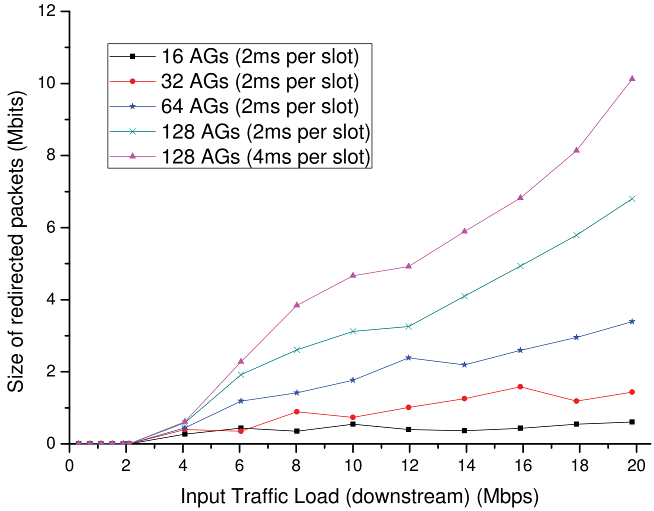

Figure 4.15: Impact of multiple AGs - size of redirected SSs

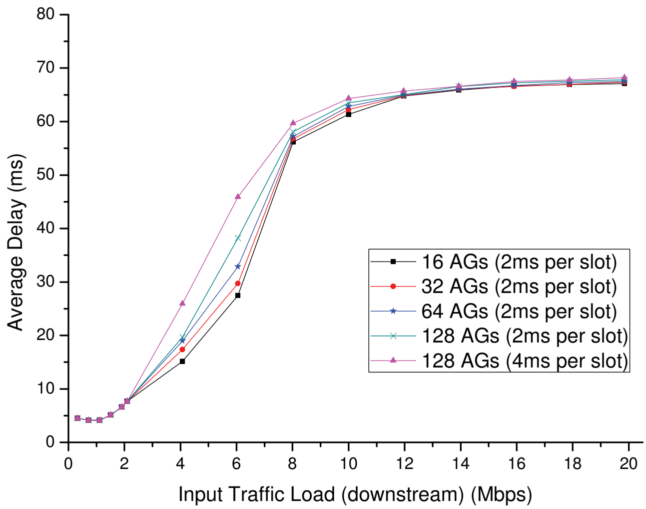

Figure 4.16: Impact of multiple AGs - average delay

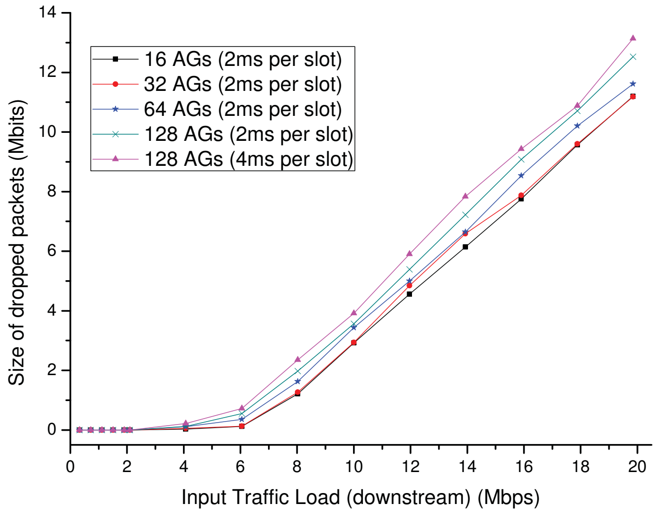

Figure 4.17: Impact of multiple AGs - size of dropped packets

reason is that the interval to update cell status becomes larger. We consider a network scenario with 64 and 128 AGs.

Figure 4.18 shows the size of redirected SSs as a function of the input traffic load of the overall network. We can see that the SPR scheme outperforms the PPR scheme in terms of resource utilization especially when the traffic load is high. The reason is that when the system load is high, the network status, such as queue size, is changed within shorter period compared to the case when system load is low. SPR scheme makes sure that the congestion in a cell can be reported and discovered on time, so that the OLT can adjust the power value of the overloaded cell and the size of redirected SSs is decreased. A relatively long interval to update cell status increases the buffer size, so that more SSs are reallocated. In other words, using SPR scheme, a cell serves more traffic and generates better network utilization.

Figure 4.19 shows the average transmission delay experienced by the data in the overloaded cell. Using PPR scheme the cell experienced high delay, because the heavily loaded situation cannot be solved on time, . On the other hand, with SPR scheme, it is observed that the

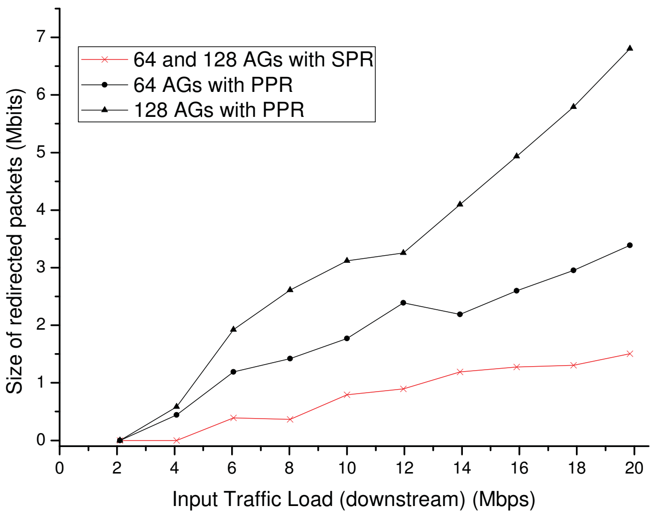

Figure 4.18: Comparison of PPR and SPR feedback schemes - size of redirected SSs

load balancing mechanism is more effective to relieve high traffic load burden from the overloaded cell.

Next, the percentage of the REPORT-p window versus the allocated upstream bandwidth is measured. The smaller percentage is, the lower overhead produced by the REPORT-p window. In this simulation, the TDM upstream bandwidth allocation is used, so that each AG is polled and assigned with fixed upstream transmission period. We define the overhead of REPORT-p window as Equation 4.19:

$$Overhead(\%) = \frac{N_{AG} \cdot (S_{basic} + N_{group} \cdot S_{addt})}{S_{BW}} \cdot 100\% \tag{4.19}$$

where N_{AG} is the total number of connected AGs. S_{basic} is the size of the basic fields in the REPORT message, including source and destination address, synchronization time, and so on. S_{addt} is the additional information, such as grouped residual queue size and residual expected transmission time, which are used for adjusting power assignment. N_{group} is the division of total cell coverage. S_{BW} is the assigned

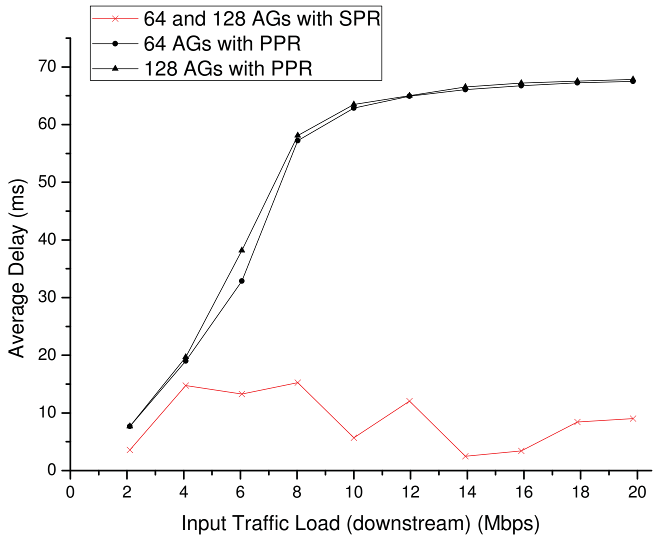

Figure 4.19: Comparison of PPR and SPR feedback schemes - average delay

upstream bandwidth to each AG, which is a fixed amount in this scenario.

Figure 4.20 shows the overhead of the REPORT-p window when the assigned upstream bandwidth for data payload is increased. We can see that when the number of AG grows, the variation of the overhead grows as well, while the overhead in the PPR case remains as the lowest. Using the PPR scheme, only the polled AG needs to transmit REPORT-p message. Therefore, the overhead is much smaller in the PPR case. We conclude that the SPR scheme improves the load balancing efficiency and results in better network performances, however, at the cost of moderately increased control message overhead.

4.6 Summary

In this chapter, an efficient load balancing mechanism using cell breathing technique has proposed for the integrated EPON and WiMAX network. A load balancing algorithm using dynamic power control to effec-

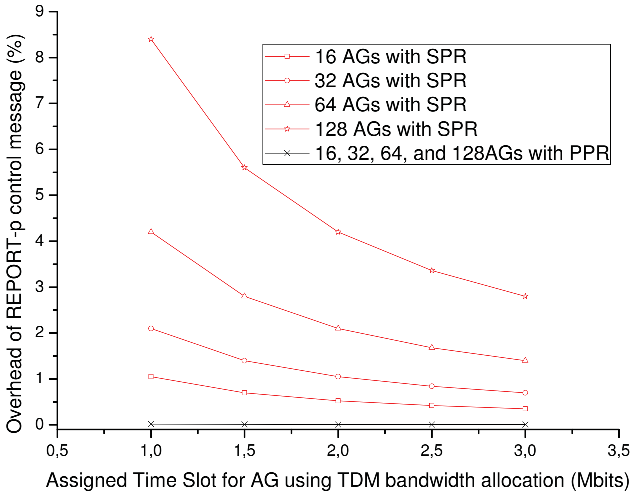

Figure 4.20: Comparison of PPR and SPR feedback schemes - overhead of REPORT-p control messages

tively adjust the size of cell coverage to maximize overall network capacity is considered. Specifically, centralized gateway selection and power allocation scheme are presented. A modified MPCP control is presented where downlink power allocations and loads feedback are made through the modified EPON control messages. In the OLT, the objective is to make jointly downlink scheduling decision and determine the transmit power. The basic idea of this algorithm is that when a cell sends out a heavily loaded indicator, the OLT starts redirecting some SSs that are located in boundary region to less loaded neighbouring cells by reducing the transmit power. The term 'cell breathing' is referring to the phenomenon where the size of a cell shrinks when the load increases, and expands when the load decreases. In this way, the heavily loaded cell can shift its load in the boundary area to its neighboring cells.

An iterative algorithm is proposed that equalizes the expected transmission time based on the reported queue sizes and adjusted transmission power level. This algorithm, refering to as the integrated load balancing problem, improves two main network performances while ensuring that users achieve QoS targets. Firstly, it reduces the possibility the case that a cell becomes overloaded, which reduces the probability of packet dropping and overdue delay. Secondly, it increases the bandwidth utilization in multi-cell system. The performance of the proposed cell breathing in integrated EPON and WiMAX system is evaluated in terms of network throughput, packet dropping probability and average queuing delay. Simulation results show that cell breathing significantly outperforms fixed power scheme.

Chapter 5

Energy Management Mechanism in Ethernet Passive Optical Networks (EPONs)

As concerns about energy consumption grow, the power consumption of the EPON becomes a matter of increasing importance. In respect of energy efficiency, the current standard has no management protocols aiming to reduce power consumption in EPONs. In this chapter, we propose an Energy Management Mechanism (EMM) for downlink EPON systems. The proposed mechanism is designed to enhance the standardized control scheme in EPON with the objective to increase energy efficiency while satisfying diverse QoS requirements. The main idea is to put an Optical Network Unit (ONU) into the sleep mode and determine a suitable wakeup time scheduler at the Optical Line Terminal (OLT). A generic EPON system is considered, which is composed of an OLT and several ONUs that are EMM enabled. An energy consumption optimization problem aimed at saving energy is proposed and two heuristic sleep mode scheduling policies are addressed to solve it. The scheduling algorithms are tightly coupled with the upstream bandwidth allocation and downstream transmission scheduling together through an integrated approach in which the total awake time in ONUs is minimized. There is a trade-off decision between maximizing the power saving and guaranteeing the network performance at the same time. Simulation results show that an EMM-based EPON with well designed scheduling disciplines is essential to achieving significant energy saving while meeting the delay and throughput constraints.

5.1 Introduction

As studied in [82–86], it has been widely recognized that reducing power consumption in data communication network becomes an important issue to global environment and future human life. Recent research has shown that, among all network segments, broadband access networks are the highest share of the network energy consumption, which consumes more than 75 percent of energy consumption by all telecommunication equipments today. Due to the rapid expansion of access network connectivity, increasing number of users and data rate are foreseeable in the future broadband access networks. As a result, energy consumption of future broadband access networks would continue to rise, with a steady annual growth rate in the near future. Therefore, it has raised attention in both academia and industry to design energy efficient network systems. Commonly, sleep mode operation is introduced to allow that nodes (or stations) can switch to sleep when they are idle and wake up when they receive or transmit packets. Such approach has been widely exploited in wireless networks, where saving battery power in mobile stations is of paramount importance due to finite power supplies in wireless devices [87–92].

The EPON, a prevailing deployed broadband access technology, provides cost efficiency and high data rate for the last mile access. A typical EPON is a Point-to-Multipoint (PMP) network with a tree based topology, where an OLT connects multiple ONUs via optical links. The OLT plays a role of distributor, arbitrator and aggregator of traffic. In the upstream direction (from ONUs to the OLT), multiple ONUs share a single link and traffic may collide. The OLT distributes the fiber capacity using an upstream bandwidth arbitration mechanism to avoid collisions. In the downstream direction (from the OLT to ONUs), data frames are broadcasted to all ONUs. ONUs filter and accept data that are addressed to them. However, ONUs have to constantly listen and ex-

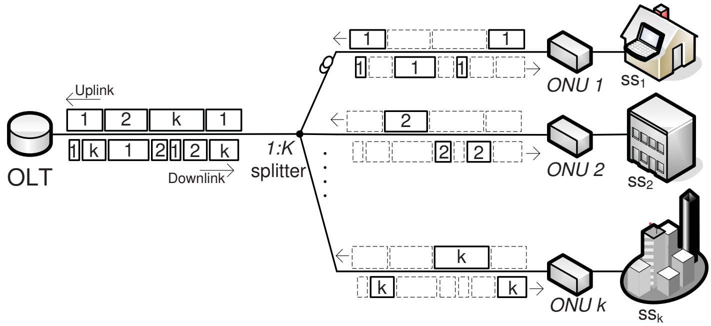

Figure 5.1: EPON upstream (multi-point to point) and downstream (point to multi-point) transmission.

amine downstream traffic, which results in wasting significant energy in the ONU. Shown in Figure 5.1, in the traditional EPON, ONUs consume energy to keep active when there is neither upstream nor downstream traffic (illustrated as the dashed boxes). As a result, minimizing power consumption is a major factor driving the design of EPON devices and of the protocols therein. An effective Energy Management Mechanism (EMM) that schedules the sleep mode period to ONUs is a key to conserve power.

The major sources of energy waste of shared resource EPON system include: overhearing, control packet overhead, and idle listening. Overhearing represents that an ONU receives and decodes packets that are not destined to it. Control packet overhead means the transmission of control messages, which are necessary to coordinate transmission between the OLT and ONUs, and unfortunately increases time and energy consumption. At last, idle listening means active time and energy ONUs spent on waiting for the next upstream and downstream transmission. It is clear that an energy efficient scheduler is designed aiming to eliminate these sources of useless energy consumption.

To reduce power consumption, ONUs are designed to enter sleep mode when they do not need to either receive or send traffic. During a sleep period, ONUs turn off the transceiver in order to save energy. In case there is incoming data for a sleeping ONU, data is queued in buffers at the OLT. Obviously, it is better to keep an ONU to stay in sleep mode as much as possible to conserve its energy. However, a power management mechanism with efficient scheduling for sleep and wake-up periods among multiple ONUs is a challenging task due to:

- ONUs should wake up and exchange traffic with the OLT. However, arrival traffic profile is not always known by the device, because the downstream and upstream traffic can either periodically or aperiodically arrive. Therefore, the sleep period should be carefully assigned in order to avoid missing any incoming packets.

- During sleep period, data is buffered and prescheduled for upstream and downstream transmission in both the OLT and ONUs. Due to QoS constraints in terms of delay and latency, the sleep period should be carefully scheduled, so target ONUs can wake up and complete transmission without violating QoS requirements.

As discussed in next Section 5.2, several studies have been proposed to analyze the power consumption for EPON, while applying the approach of allowing ONUs to sleep mode. Although different aspects have been investigated on how to minimize energy consumption in PON systems, there have been few studies focusing on the protocol design for supporting sleep mode ONUs in EPON. To implement EMM in EPON, the legacy control scheme requires modifications and extensions. Moreover, very few have explicitly addressed the trade-off between energy consumption and network performances. To the best of our knowledge, our work is the first to propose the assignment of sleep period to ONUs taking into account of the traffic transmission in both upstream and downstream directions.

The remainder of this chapter is organized as follows. We first give a brief description of related works in Section 5.2. The introduction of the system model is addressed in Section 5.3. The design of EMM based EPON with two different downstream schedulers is introduced in Section 5.4. Simulation environments and results are outlined and discussed in Section 5.5. Finally, conclusions and future work are drawn in Section 5.6.

5.2 Related Works

Issues of energy efficiency have been studied in EPONs as well as in other optical access networks. In this section, we present related works and compare their models with ours.

Authors in [93] propose a management protocol for ONUs to request entering in sleep mode. The OLT reserve a minimal bandwidth for ONUs in sleep mode. During a sleep period, no connection is established and no traffic is delivered. The sleep mode can be terminated by either the OLT or the ONU by using an administration message. This protocol eliminates conventional requirements for a complete activation procedure in ONUs. The authors address the impact of sleep mode protocol for re-activation process in GPON.

In [94], the authors present implementation challenges of sleep mode operation in PONs, such as achieving clock synchronization and reducing the clock recovery overhead. The process of regaining network synchronization after waking up from sleep mode is studied in both GPON and EPON system. Two novel ONU architectures are proposed and compared with current ONU architectures on power consumption and wakeup overhead. As already indicated, we focus on protocol design and do not address the physical implementation issues in our work. The architectures as well as the overhead values proposed in [94] can be used for this purpose.

Another paper [95] proposes a power saving mechanism for 10 Gigabit class PON, where ONUs are implemented with a sleep and periodic wakeup regime. The sleep mode is triggered based on negotiations between the OLT and the ONU, and a variable sleep time is allowed. The authors use a prediction model of traffic profiles to determine the initiation and the duration of sleep mode. A probabilistic analysis of the presence or absence of downstream traffic is required in order to ensure the ONU can wake up and the packet is delivered on time. This model differs from ours in that our model does not use estimation and the energy management mechanism is mapped in EPON MAC layer.

Author in [96] presents the EPON power saving issue on the IEEE 802.3az meeting in September 2008. The functions for power saving are proposed to be implemented in both the OLT and ONUs. The initiation and termination of sleep mode by using handshake messages are drafted.

The attention of energy efficiency is raised in the IEEE standardization, however, there are considerably few research works in saving energy consumption of optical network management systems. Our work investigates the current control mechanism for EPON and proposes control protocol for sleep mode operation.

The relationship between energy and data rate in general optical IP networks was also studied in a different context. For example, [97] presents power saving of different network architectures, including all-optical networks, first-generation networks and multi-hop networks. Power consumption of electronic versus optical components in optical WDM networks is investigated under different traffic loads. In [85] this problem is generalized by considering the power consumption in an optical IP network. A new network-based model is established, which includes core, metro and access networks.

5.3 System Model

In this chapter, we consider an EPON system consisting of an OLT, $1:K$ splitter and multiple ONUs. The upstream and downstream traffic are separated in different wavelengths, typically 1310 nm for the upstream transmission and 1550 nm for the downstream transmission. TDMA is used in the physical layer where bandwidth is divided in time slots. Each ONU maintains an upstream buffer and sends upstream data to the OLT during assigned time slots. For simplicity, we assume in this chapter that the upstream data rate is fixed for all ONUs. In the downstream direction, when the arrival rate at the OLT exceeds the output data rate, data are queued in the OLT and transmitted when downstream bandwidth is available.

We also assume that the OLT and ONUs can regain synchronization successfully so that ONUs can enter sleep mode and return back to awake mode. The power consumption value in the awake state includes energy consumed in listening to the OLT and in receiving or transmitting data. We neglect the transition delays and the energy consumption when an ONU changes states. We assume that the packets are able to be served within an integer number of time slots. During the observation period, the number of ONUs is also fixed. Next we list some notations (Table 5.1) used in the rest of this work and then formulate the energy

Notations	**Value**
K	The number of ONUs registered in the OLT.
i	The index of ONUs in EPON. $i \in K$
N	The total time slots within the overall observed transmission period.
n	The index of time slot. $n \in N$
s_i^n	ONU i state indication (awake mode or sleep mode) in time slot n.
T_i^n	Period of time slot n in ONU i.
P_{awake}	The average amount of energy (Watt) consumed during the awake period.
P_{sleep}	The average amount of energy (Watt) consumed during the sleep period.

Table 5.1: Notations in energy efficiency mechanism

saving scheduling problem.

Assuming that the overall transmission period is divided into N time slots, ONUs are either in awake mode or in sleep mode during the time slot n. The total energy expenditure model is formalized as follows:

$$\mathbf{E} = \sum_{i=1}^{K} \sum_{n=1}^{N} T_i^n \cdot [s_i^n \cdot P_{\text{awake}} + (1 - s_i^n) \cdot P_{\text{sleep}}] \tag{5.1}$$

5.3.1 Analysis of Sleep Mode Operations

Within a period, the total energy consumption includes energy consumed by M ONUs, which are in active state. As expressed in our energy expenditure model (Equation 5.1), the total energy consumed by each ONU consists of two parts: in either the *awake* or *sleep* mode. In the awake mode, there are three different states: receiving packets, sending packets and idle listening. The idle state is defined as the time when an ONU is active without any receiving or transmitting action. In the sleep state, ONUs switch off transceiver functions and keep a timer on to count down the wakeup time. Nodes consume much less power in the sleep mode than in the awake mode. For instance, regarding to different architectures for a sleep mode enabled ONU [85], the expected power consumptions require 2.85 W for active and 750 mW - 1.28 W for sleeping. When ONUs enter sleep state, ONUs are in the absence of traffic at the optical transmission interface. However, some functions, such as clock and data recovery and back-end digital circuit, are still powered on. In this sense, the sleep mode differs from power down, which is usually an indication of any equipment fault or disconnection in EPON. If an ONU is reported as deactivation after a power down, the ONU has to register itself to the OLT again via the periodic discovery processing. On the other hand, using the sleeping mode, ONUs can wake up, listen to, and communicate with the OLT without the registration process, when the sleep period is finished.

As shown in Figure 5.2, a timer is utilized in the sleep mode control. Upon receiving notification of entering the sleep mode, the timer counts down and the ONU wakes up when the timer is expired. When an ONU wakes up from the sleep mode, there is an overhead time window (T_{ov}) for recovering the OLT clock (T_{recv}) and retrieving network synchro-

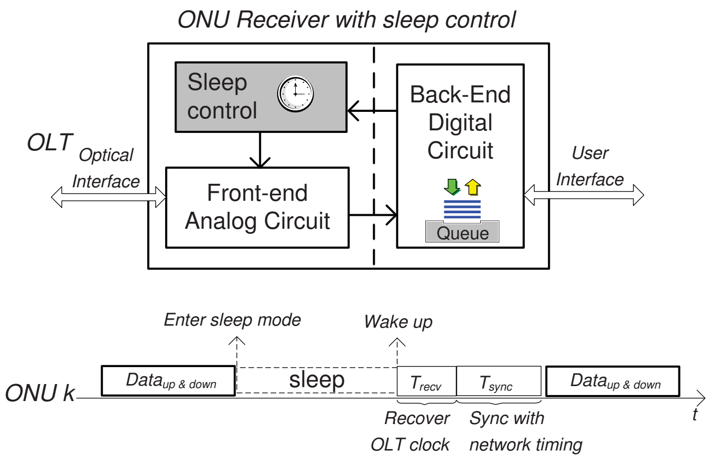

Figure 5.2: ONU receiver with sleep control function.

nization (T_{sync}). The clock recovery time varies in different implementations of ONU receiver architectures. For synchronization process, a fixed Start Position Delimiter (SPD) is embedded in the EPON control message header, which allows ONU to gain synchronization with the start of the EPON frame. The synchronization time in EPON is up to 125 μs [94].

Following our model, it is obvious that energy depletion is determined by the time the ONU spends in awake mode. An optimal EMM solution is to schedule ONUs into sleep mode whenever there is no transmission in order to minimize the overall power consumption. However, since there are QoS requirements in terms of delay and throughput, it requires a scheduling policy to improve energy efficiency without violating the QoS requirements. Therefore, this work deals with the trade-off between the allocation of wakeup frequency and the queuing delay in the EPON.

5.3.2 Analysis of Energy Consumption and Delay Performance

When energy management mechanism is implemented in EPON, ONUs turn their optical communication interface on and off to minimize energy consumption. Therefore, in order for an ONU to communicate with the OLT, it must be in active mode. As a centralized control station, the OLT schedules active duty cycles and informs connected ONUs about its assignments. An ONU must comply with the noticed sleep period and wakeup time. When an ONU is in sleep mode, the downstream traffic destined to the ONU is queued in the OLT. This communication model imposes a clear trade-off between the delay encountered by a packet and the time during which ONUs are in active mode. Solutions for addressing this trade-off depend, to a large extent, on the following aspects:

1. The expected amount of data to be received and transmitted.

2. The polling scheme and upstream bandwidth allocation scheme: for example, the sequence of polled ONU and the size of bandwidth granted to the ONU.

3. Whether the OLT process the packets and send to ONUs with priority discrimination.

With respect to 1, if the data arrival rate is high and continuously delivered, data are queued and the OLT can schedule transmission. With respect to 2, if the polling scheme and upstream bandwidth allocation are flexible enough, it can determine the order of ONUs to communication according to the expected delay constraint. With respect to 3, if the OLT can process packets, it may buffer and merge several packets for aggregation.

We first formulate our scheduling problem: find an optimal scheduling discipline minimizing the average energy consumption of EPONs while guaranteeing an upper bound on the queuing delay. This problem is formalized as an optimization problem. An optimal assignment of wakeup time to ONUs is an assignment that guarantees an upper bound D_{max} on the maximum delay while minimizing the total energy spent by the ONUs in awake mode. Let K be a set of ONUs that are awake and communicate with the OLT during the whole N observation periods. Let T_i^n and P_{awake} be the awake period and the average energy consumption for each ONU i, respectively. Then, the goal is to:

$$minimize: \quad E = \frac{1}{K}\sum_{i=1}^{K}\sum_{j=1}^{N} T_i^j \cdot P_{awake} \tag{5.2}$$

subject to:

$$d_i^m < D_{max} \tag{5.3}$$

where d_i^j is the delay experienced by every connections in AG i, which should deliver qualified services.

5.4 Energy Management Mechanism (EMM)

In this section, we introduce the implementation of EMM together with EPON MAC protocol, Multipoint Control Protocol (MPCP) [36]. The OLT is in full control of the bandwidth allocation in both downstream and upstream. An ONU can associate with the OLT either in normal mode, referred to as Power Ignoring Scheme (PIS), or in EMM mode. When an ONU is in normal mode, it constantly stays awake and consumes power, i.e., never goes into sleep mode. In the following, we assume that all ONUs are EMM-enabled.

In order to minimize power consumption, ONUs remain in a sleep mode most of the time while adhering to the assignment from the OLT. We show how being in EMM yields considerable power saving for an ONU as compared to PIS case. In addition, we introduce how the sleep mode scheduling scheme has an impact on the energy and delay performance of the EMM ONUs. The sleep period and wakeup time can be calculated and assigned using either an Upstream Centric Scheduling (UCS) algorithm or a Downstream Centric Scheduling (DCS) algorithm.

5.4.1 Energy Efficient Scheduler Design

In EPONs, the downstream and the upstream transmission are separated, and the OLT plays a central role to reserve bandwidth and to serve traffic transmission. In this chapter, a *subframe period* (SP), T_{sp}, refers to the time assigned to an ONU to send upstream data to the

OLT (SP-UL) or to receive downstream data from the OLT (SP-DL). In this work, we will analyze the energy consumption and determine the wakeup time by considering both the upstream and downstream data since the instants of terminating sleep mode by upstream and downstream transmission are different.

- ***Downstream bandwidth allocation***: In the MPCP specified in IEEE 802.3ah, the OLT broadcasts a downstream packet with destination MAC address to all ONUs. ONUs that are awake to listen if the OLT has packets to deliver to them check the header for destination identification. Since there is no retransmission mechanism in the EPON downstream, it takes a long latency if the ONU misses packets. Therefore, ONUs keep active all the time in order to listen to the OLT and receive packets if the destination identification is matched. Since the power saving mechanism is ignored traditionally, time and energy are wasted when an ONU is active to listening to the OLT and decoding packets destined to other ONUs.

- ***Upstream bandwidth allocation*** In the MPCP, the OLT maintains a polling list and a polled ONU has the right to transmit data within its assigned upstream time slots. The OLT allocates upstream time slots to each ONU in order to avoid that multiple ONUs transmit to the uplink simultaneously. The upstream bandwidth information such as the start time and the length of granted transmission window is carried in the GATE message. The OLT polls each ONU by sending them GATE message and ONUs transmits traffic within the assigned upstream time slots after receiving the GATE message. The upstream bandwidth is distributed among all ONUs by using either a fixed allocation scheme or a dynamic bandwidth allocation (DBA) scheme [37]. In the former case, each ONU is scheduled with a fixed amount of upstream transmission period. In the latter case, the bandwidth is allocated dynamically with a bounded size according to ONUs' request. A cycle (T_{cycle}) refers to an interval between two successive polling messages to an ONU. Since the power saving mechanism is ignored traditionally, the time is wasted when an ONU is active to waiting for its next upstream transmission period.

The EPON supports energy management at the MAC layer, and ONUs need to negotiate with the OLT in order to decide the power saving parameters. The power saving parameters include the time to sleep and wake up. Theoretically, the power saving control can be initiated either at the OLT or at the ONUs, if they are in an idle state, where no packet to be received or sent. However, as the ONU is an aggregation node that collects traffic from low level networks and relays to the OLT, it is commonly assumed that ONUs always have queued data for upstream transmission. In this work, we consider only the OLT to trigger the power saving mechanism. We focus on protocol design of EMM and the scheduling schemes to maximize power saving subject to the constraint of packet delay.

With respect to bandwidth allocating, since the MPCP is central control protocol, the OLT has full knowledge to assign transmission bandwidth in both upstream and downstream directions. This implies that the order of granted ONUs and the amount of granted bandwidth are determined by the OLT. An intuitive solution for EMM scheduler is to aggregate the downstream traffic in order to minimize the time allocated to ONUs for idle states. The general design of EMM based EPON proposes following functions in the OLT and ONUs:

1. ***OLT operation***:
   The OLT can buffer downstream traffic and schedule the appropriate transmission period to each ONU. To meet the QoS requirement of delay sensitive traffic, the downstream bandwidth is allocated such that every packet can be served before its deadline. The original control message GATE is modified with additional fields indicating the assigned sleep time and wakeup time. To assign the wakeup time for the next upstream transmission, the OLT needs to calculate the time for an ONU to upload and download data packets. Considering the wakeup time to transmit the upstream packet, the OLT determines and allocates the upstream transmission windows for all ONUs. As for the wakeup time to receive the next downstream packet, since the number of buffered downstream data is completely known by the OLT, it can calculate and schedule downstream transmission.

2. ***ONU operation***:

After scheduling the sleep period, the OLT sends a control message to the ONU for the permission to transit into sleep mode. Upon receiving this message with parameters start time for sleep and wakeup time from sleep, the ONU enters into sleep mode. After a sleep mode, the ONU transits back to the awake mode. Scheduled ONUs should wake up according to the assigned wakeup time and check the sleep-mode GATE message (G_s). Upon receiving the G_s message, the ONU derives the sleep period and obeys the assigned upstream and downstream bandwidth allocation. As for the upstream subframe period announcement, the assigned wakeup time implies its next access time to the upstream transmission. For a G_s message announcing the downstream subframe period, the ONU should awake at the notified wakeup time to receive the buffered data and then back to sleep afterwards.

5.4.2 Upstream Centric Scheduling (UCS) Algorithm

The main idea of UCS scheme is that the OLT assigns the awake period to ONUs according to their corresponding upstream allocation. The OLT grants the upstream bandwidth by polling each ONU. During the granted upstream subframe period, the ONU is awake. Once an ONU transits into the awake state, the OLT with UCS algorithm only transmits downstream packets to the awake ONU, and queues those packets destined for 'sleep' ONUs into a buffer.

In this subsection, we give a description of the UCS algorithm. For simplicity of illustration, we consider a system of an OLT and three ONUs for two upstream transmission periods. For more ONUs and transmission periods, the same logic can be applied. Figure 5.3 shows the process of UCS protocol. In this work, the t denotes for the instant time and T represents for a period. The OLT maintains an entry table, which contains 1) the start transmission time (${t_{start}}_i^n$) and granted bandwidth (BW_i^n) for allocating upstream bandwidth; 2) the sleep time (${t_{sleep}}_i^n$) and the wakeup time (${t_{wake}}_i^n$) for assigning sleep period. The table is updated every cycle n for ONU_i. Downlink data are stored in subqueues for each ONU.

As shown in Figure 5.3, the ONU_1 wakes up at t_0 and receives sleep mode GATE message (${G_s}_1^1$) at the time t_1. After learning the granted

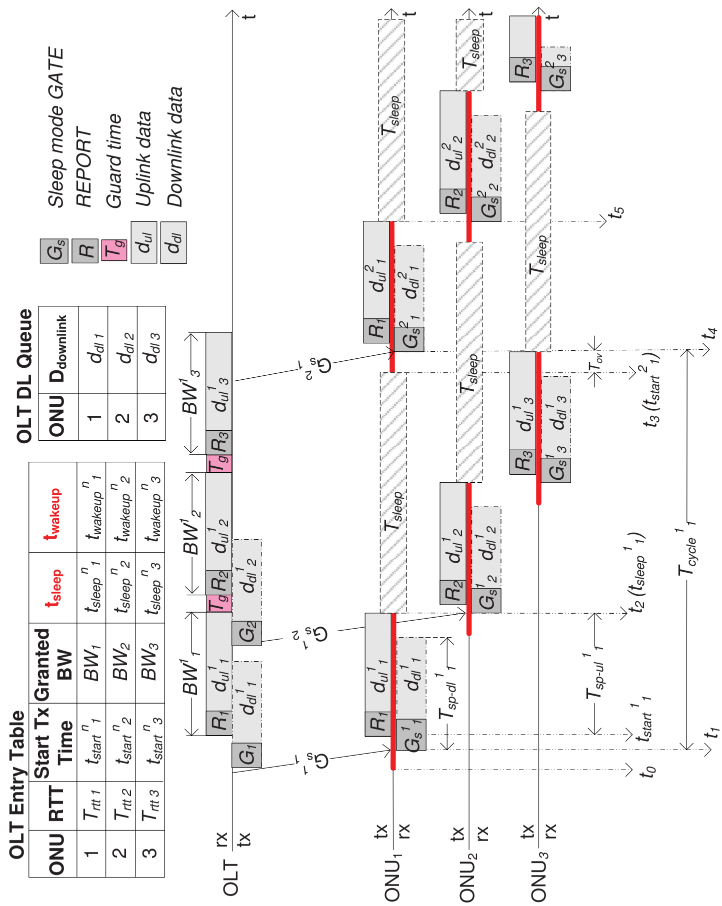

Figure 5.3: Sleep mode operation with Upstream Centric Scheduling (UCS) scheme.

upstream subframe period (${T_{sp-ul}}_1^1$), ONU_1 generates REPORT message (R_1) and starts upstream transmission (${d_{ul}}_1^1$) complying with the assignment. During ${T_{sp-ul}}_1^1$, the OLT derives downstream data (${d_{dl}}_1^1$) from the subqueue for the ONU_1 and transmits them to the ONU_1. Notice that the data d_{dl} represents either one data packet or a series of integrated data packets. The start time of sleep mode (${t_{sleep}}_1^1$) for ONU_1 is assigned so that the ONU_1 enters into sleep mode immediately when the upstream subframe period is completed at t_2. The calculations of upstream transmission period and the time to turn into sleep mode are listed in Equation 5.4 and Equation 5.5.

$$ {T_{sp-ul}}_i^n = \frac{BW_i^n}{R_o}, \quad i \in K, n \in N \tag{5.4} $$

$$ {t_{sleep}}_i^n = {t_{start}}_i^n + {T_{sp-ul}}_i^n, \quad i \in K, n \in N \tag{5.5} $$

where BW_i^n is the allocated upstream bandwidth for the i^{th} ONU during the n^{th} cycle. R_o is the transmission rate of the optical upstream and T_g is the guard time between two successive upstream transmissions. The wakeup time (${t_{wakeup}}_1^1$), which is same as the length of sleep period, is calculated based on the period before the ONU_1 is polled again (${T_{cycle}}_1^1$). In the Figure 5.3, the next wakeup time for the ONU_1 is t_3. The value of wakeup time is computed at the OLT and assigned to each ONU. The polling cycle is computed using Equation 5.6.

$$ \begin{aligned} {T_{cycle}}_i^n &= \sum_{i=1}^{K} ({T_{sp-ul}}_i^n + T_g) \\ &= \sum_{i=1}^{K} (\frac{BW_i^n}{R_o} + T_g), \quad i \in K, n \in N \end{aligned} \tag{5.6} $$

After a period of T_{cycle}, the interval between two adjacent polling messages, an ONU is polled again at t_4 (t_1 + ${T_{cycle}}_1^1$). The calculation of ONU wakeup time needs to take the overhead period into account due to the time of recovering the OLT clock and retrieving the network synchronization. The computation of the overhead period (T_{ov}) and the time to wake up are listed as Equation 5.7 and Equation 5.8, respectively.

$$T_{ov} = T_{recv} + T_{sync} \tag{5.7}$$

$$t_{wakeup}{}_i^n = t_{start}{}_i^n + T_{cycle}{}_i^n - T_{ov}, \quad i \in K, n \in N \tag{5.8}$$

During the observation period, the total awake time for each ONU is the sum of allocated upstream time slots and an ONU enters the sleep mode otherwise. During the subframe period, the ONU_i transmits upstream data $d_{ul}{}_i^n$ and receives downstream data $d_{dl}{}_i^n$, within the granted time slot n. Notice that the size of the uploaded and downloaded data must equal or less than the upstream bandwidth, i.e. size($d_{ul}{}_i^n$)$\leq BW_i^n$ and size($d_{dl}{}_i^n$)$\leq BW_i^n$. The total energy consumed by the ONU i is illustrated in Equation 5.9:

$$\begin{aligned} \boldsymbol{E}_{\mathbf{UCS}} = & \sum_{i=1}^{K}\sum_{n=1}^{N}[T_{sp-ul}{}_i^n + T_{ov}] \cdot P_{awake} \\ & + \sum_{i=1}^{K}\sum_{n=1}^{N}[T_{cycle}{}_i^n - T_{sp-ul}{}_i^n - T_{ov}] \cdot P_{sleep} \end{aligned} \tag{5.9}$$

The UCS based scheduler is simple because the sleep period is determined based on the upstream transmission. An OLT with the UCS based scheduler uses the awake ONU, which is assigned to the upstream transmission as destination to retrieve buffered data. The algorithm is presented in the flowchart in Figure 5.4. However, the performance of downstream data latency and bandwidth utilization may not be satisfied due to the dependency on the upstream subframe period, the upstream polling sequence, and the total number of active ONUs. One disadvantage of this scheme is that it is not suitable for delay-sensitive traffic due to longer queuing delay experienced in the OLT.

5.4.3 Downstream Centric Scheduling (DCS) Algorithm

The second scheduling policy presented in this chapter is that the OLT stores downstream traffic in a First-In First-Out (FIFO) buffer and the ONU has to be awake and receive downstream traffic whenever the OLT sends one. The awake periods are assigned to favor both upstream and downstream traffic. The DCS scheme is illustrated in Figure 5.5. The

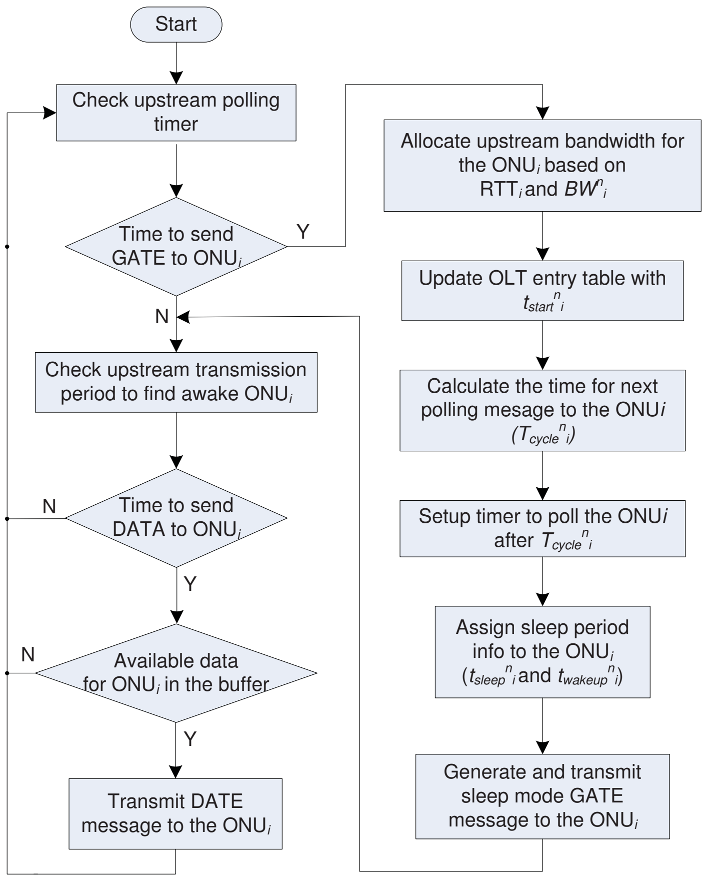

Figure 5.4: Flowchart of the UCS algorithm designed in the OLT.

sleep mode GATE message, ${G_s}_i^n$, is used for both upstream bandwidth allocation and sleep period assignment. The upstream transmission period is calculated using the same equations 5.4- 5.8 as in the UCS scheme. For example, the ${G_s}_1^1$ at time t_1 indicates that the next polling time is at time t_8 and the ONU_1 needs to wake up at time t_7. In the downstream, the OLT serves data in a first-in first-out order. For instance, shown in Figure 5.5, during the upstream subframe, ${d_{ul}}_1^1$, the OLT sends packets to both ONU_1 (data ${d_{dl}}_1^1$) and ONU_2 (data ${d_{dl}}_2^1$).

As shown in Figure 5.5, the admission information for both upstream and downstream transmission is notified by the sleep mode GATE message. The ${G_s}_1^1$ informs ONU_1 the allocated upstream transmission of ${d_{ul}}_1^1$ and assigned sleep period during t_3 to t_7. After transmitting ${d_{dl}}_2^1$ to the ONU_2, the OLT gets data ${d_{dl}}_1^1$ from the head of its downlink queue and sends to the ONU_1 together with ${G_s}_1^2$. Since there is data ${d_{dl}}_1^2$ in the buffer, the ONU_1 is notified to wake up at t_4 instead of t_7. The sleep period is assigned if there is neither upstream nor downstream transmission scheduled. If the OLT has queued data for the ONU, the wakeup time for the next data can be calculated. For example, the period between t_3 and t_4 is the idle period for ONU_1. Otherwise, the OLT sends a message to inform the ONU of remaining in awake mode in order to avoid missing any downstream traffic. In the following, the assignment of sleep period is discussed in details.

Under the condition that the downstream traffic terminates the sleep mode and the OLT assign the sleep period for the i^{th} ONU, we distinguish four possibilities as shown in Figure 5.6. The sleep period is determined by favoring both the successful reception of upstream and downstream data. Accordingly, we denote G_s as the GATE control message carrying the information of sleep period, such as the start time of sleep period (t_{sleep}) and the wakeup time (t_{wakeup}). Let ${T_{sp-ul}}_i^n$ and ${T_{sp-dl}}_i^n$ be the n^{th} subframe period of uplink and downlink transmission, respectively. In addition, the T_{ov} period represents an overhead time for clock recovery and synchronization.

1. Case0: GATE0 is used to inform the allocated upstream bandwidth for the ONU_i. Because the OLT has full knowledge about the upstream bandwidth allocation, the sleep interval is determined, including both the sleep time (t_{sleep0}) and the wakeup time ($t_{wakeup0}$).

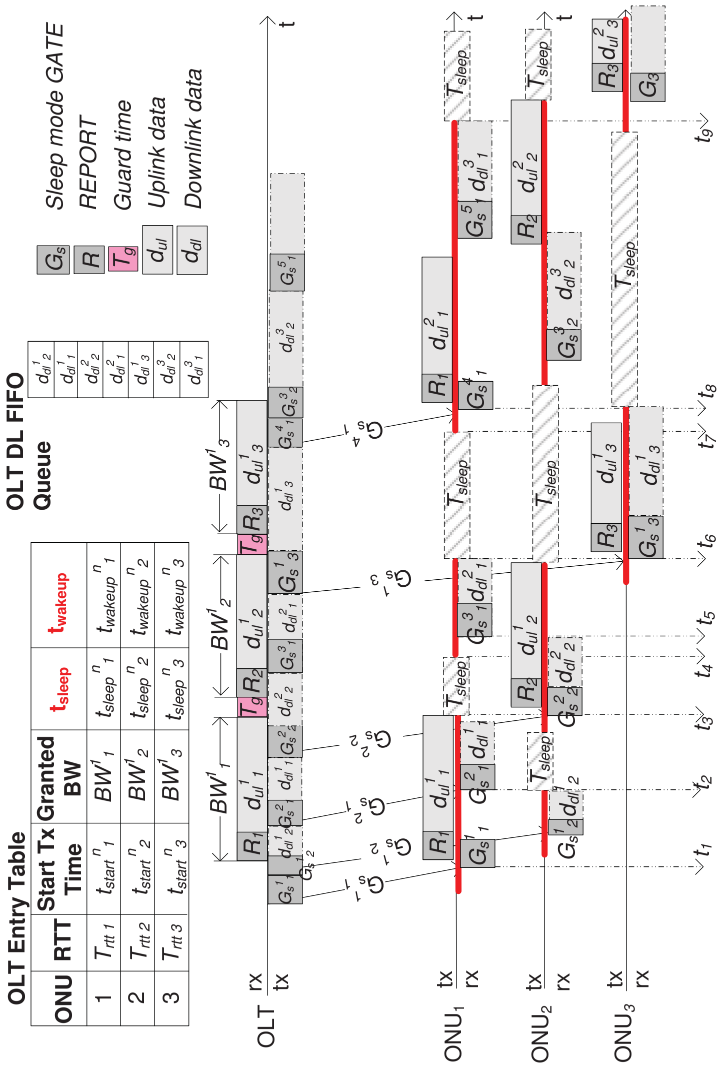

Figure 5.5: Sleep mode operation with Downstream Centric Scheduling (DCS) scheme.

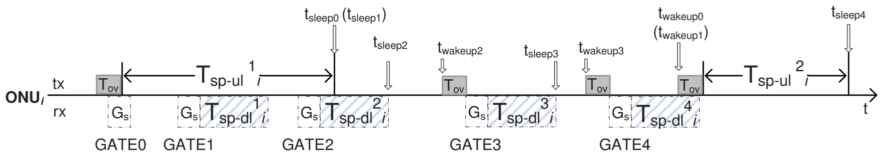

Figure 5.6: Downstream transmission in the sleep period model.

2. Case1: The transmission of downstream data ($T_{sp-dl}\,{}^{1}_{i}$) is completed within the allocated upstream subframe period. If there is other queued data in the OLT, such as the payload with GATE2, the sleep period is assigned as same as in Case0. If the next received data is the payload with GATE3, the time to wake up is reassigned as $t_{wakeup2}$. If there is no more queued data in the OLT for the ONU_i, the original assignment, t_{sleep0}, is removed, because ONU_i should keep awake in order to avoid missing future arrival data.

3. Case2: In this case, the downstream data ($T_{sp-dl}\,{}^{2}_{i}$) cannot be finished before the start time of the sleep period as assigned in Case0 and Case1. Therefore, the start time of sleep period is postponed to t_{sleep2}. The OLT examines its downstream queue and schedules the transmission of the next downstream data ($T_{sp-dl}\,{}^{3}_{i}$) for the i^{th} ONU. Moreover, the wakeup time ($t_{wakeup2}$) is calculated in order to ensure a successful transmission.

4. Case3: In this case, downstream data is received during the sleep interval. In GATE3 control message, the new time of entering the sleep mode is updated, t_{sleep3}, which is calculated based on the downstream subframe period ($T_{sp-dl}\,{}^{3}_{i}$).

5. Case4: As shown in the last GATE4 in this figure, downstream data is arrived during the sleep interval and lasted till the next upstream subframe period. Until the sleeping period is completed, the ONU transits into the awake mode and will receive the GATE4 message with the information of upstream bandwidth allocation, such as the start time and length of $T_{sp-dl}\,{}^{4}_{i}$. Thus, the GATE4

message contains bandwidth assignments for both upstream and downstream subframe period.

The flowchart of the DCS algorithm implemented in the OLT is described in Figure 5.7. Compared to the UCS algorithm, the OLT checks the available data in the buffer and update the sleep period for the destination ONU with a sleep mode GATE message. The wakeup time is precisely calculated and assigned, so that the ONU can be active to carry out the next upstream or downstream transmission. Since the operations of the above scheduling mechanism requires the OLT to locate the next packet in the buffer in order to calculate the wakeup time, ONUs which have no more queued packets cannot enter the sleep mode. In this case, the OLT assigns the sleep period as zero.

In order to calculate the total energy consumption in the DCS algorithm based EMM, we define the awake period for upstream transmissions ($T_{awake-ul}$) and downstream transmissions ($T_{awake-dl}$) separately. First, we compute the awake period and sleep period in the upstream direction (shown in Equation 5.10 and Equation 5.11), which is similar to Section 5.4.2.

$$T_{awake-ul}{}_i^n = T_{sp-ul}{}_i^n + T_{ov}, \quad i \in K, n \in N \tag{5.10}$$

$$T_{sleep-ul}{}_i^n = T_{cycle}{}_i^n - T_{sp-ul}{}_i^n - T_{ov}, \quad i \in K, n \in N \tag{5.11}$$

Next, we assume that there are M downstream transmissions to the i^{th} ONU. For the downstream, the awake period is calculated (in Equation 5.12) differently in each case mentioned in Figure 5.6. In Case1, ONU_i is in awake state, so that the awake period for receiving downstream data is zero. From Case2 to Case4, if the time for waking up and sleeping is assigned, the period that the ONU has to wake up from its sleep mode is calculated as the interval between t_{sleep} and t_{wakeup}. Illustrated in Equation 5.13, the sleep period is the upstream sleep period minus the awake period used for the downstream transmission, which occurs during the upstream sleep period.

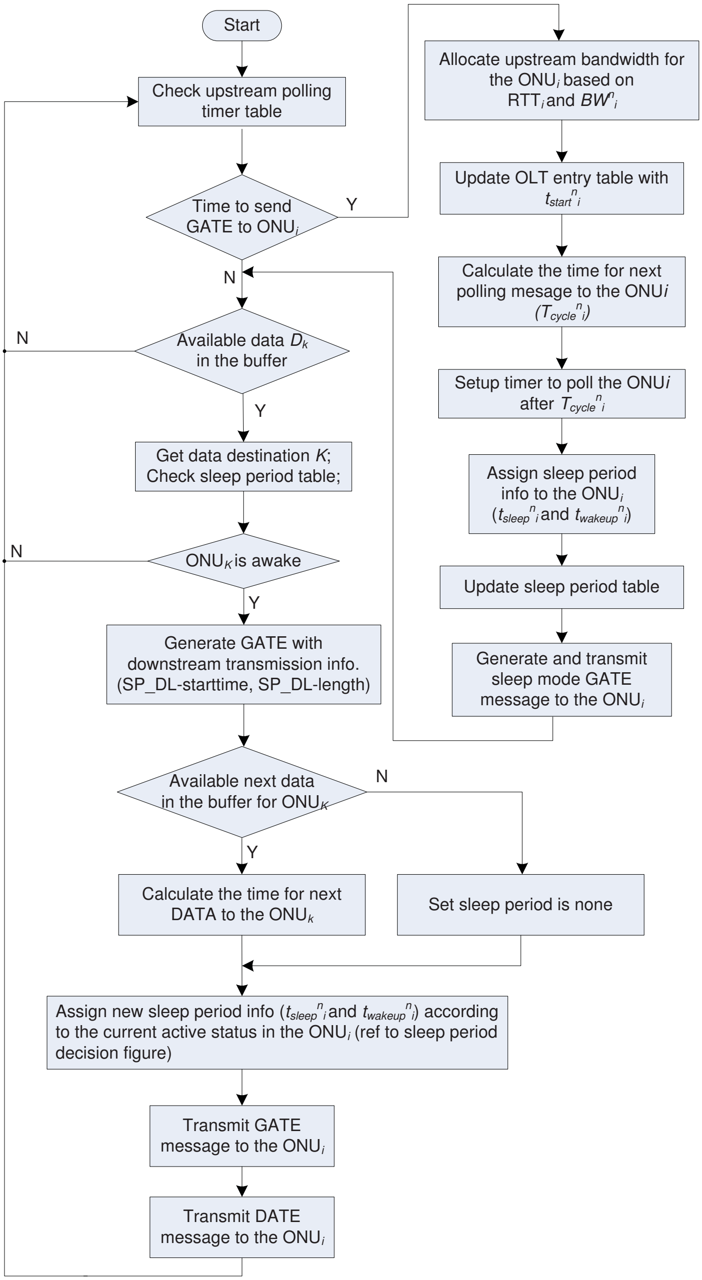

Figure 5.7: Downstream transmission in the sleep period model.

$$T_{awake-dl}{}_i^n = \sum_{m=1}^{M} T_{awake-dl}{}_i^m$$
$$= \sum_{m=1}^{M} \begin{cases} 0 & \text{, if Case1} \\ t_{sleep2}^m - t_{sleep1}^m & \text{, if Case2} \\ t_{sleep3}^m - t_{sleep2}^m & \text{, if Case3} \\ t_{sleep0}^m - t_{sleep3}^m & \text{, if Case4} \end{cases} \quad i \in K, m \in M, n \in N \tag{5.12}$$

$$T_{sleep-dl}{}_i^n = T_{sp-ul}{}_i^n - \sum_{m=1}^{M} T_{awake-dl}{}_i^m, \quad i \in K, n \in N \tag{5.13}$$

After analyzing the active behavior of an ONU in both the upstream and downstream transmissions, the total awake and sleep periods are concluded as Equation 5.14 and Equation 5.15:

$$T_{awake-total}{}_i^n = T_{awake-ul}{}_i^n + T_{awake-dl}{}_i^n, \quad i \in K, n \in N \tag{5.14}$$

$$T_{sleep-total}{}_i^n = T_{sleep-dl}{}_i^n, \quad i \in K, n \in N \tag{5.15}$$

The total energy consumed is calculated in Equation 5.16. The DCS sleep mode scheduler transmits downstream traffic in a flexible way, unlike in the UCS that only the active ONU scheduled with an upstream transmission period can receive data from the OLT. The OLT tries to schedule the queued downstream traffic in a way to minimize the delay. The UCS is simple where the sleep period and wakeup time is determined only by the upstream bandwidth allocation. On the other hand, the DCS scheduler requires to keeping track of both downstream and upstream transmission windows.

$$\begin{aligned} \boldsymbol{E}_{\mathbf{DCS}} &= \sum_{i=1}^{K} \sum_{n=1}^{N} [T_{sp-ul}{}_i^n \cdot P_{awake}] \\ &+ \sum_{i=1}^{K} \sum_{n=1}^{N} [T_{sleep-total}{}_i^n \cdot P_{sleep}] \end{aligned} \tag{5.16}$$

5.5 Simulation Results

5.5.1 Simulation Scenario

In this section, we study performances and features of our proposed energy management mechanism compared to the fixed power control, which is represented as Power Ignoring Scheme (PIS). Simulations have been carried out by means of OPNET network simulator [60] to measure the performance metrics. The performed simulations aimed to show the diverse effects of system parameters, such as the upstream scheduling scheme, the number of connected ONUs, and the arrival packet size. Then the benefits that can be obtained while utilizing the proposed EMM in EPON. The following parameters are monitored, such as the power consumption, the network throughput, and the average queuing delay. The average awake time in ONUs is used as a merit for energy efficiency performances. The less time an ONU is awake, the less energy an ONU consumes.

We have assumed that the OLT is connected to K ONUs (K=16 or 32)., which are enabled with power saving functionality. The optical link rate is 1 Gbps for both upstream and downstream transmission. The guard time between two consecutive transmission slots is 5 us. For allocating upstream bandwidth among ONUs, both fixed scheme (e.g. TDM) and dynamic scheme (e.g. IPACT) are implemented at the OLT. The maximum transmission cycle is set to 2 ms. As the proposed power saving mechanism with two scheduling schemes cooperates the optical downstream bandwidth scheduler at the OLT, it is reasonable to monitor the performance of downstream traffic. Traffic is generated following Poisson distribution and the length of packets is generated independently between 64 bytes to 1200 bytes. Destinations of downstream traffic are uniformly distributed among connected K ONUs. The MPCP control messages, GATE and REPORT, are formatted according to the IEEE 802.3ah specification. The sleep mode GATE message, G_s, is an extension of the standard GATE message. The additional two 4 bytes fields are embedded as assigned sleep time and wakeup time. The total size of G_s message is set to 41 bytes.

5.5.2 Validation of EMM with the fixed upstream bandwidth allocation

In Section 5.4, we analyze the energy efficiency mechanism with two proposed scheduling approaches and derive the consumed power accordingly. The simulation scenario consists of 16 ONUs and the OLT grants fixed upstream bandwidth to each ONU. The results in Figure 5.8 refer to the system behavior in the presence of an fixed upstream bandwidth allocation scheme deployed at the OLT. They show the energy consumption, represented by the ONU awake time, when the downstream traffic load changes. In the case of PIS, ONUs constantly stay in awake state and consumes high energy (100% of the observation period). When the EMM is enabled in the EPON system, the awake time maintains at a low average percentage value if the Uplink Centric Scheduling (UCS) scheme is applied. As explained in Section 5.4.2, the period of being power on is decided by allocated upstream transmission periods in the UCS case. When each ONU is allocated with fixed amount of upstream bandwidth, the awake time is also fixed. On the other hand, under the DCS case, ONUs wake up and communicate with the OLT for either upstream or downstream traffic. When arrival traffic is less than the optical output link rate, the awake time of EMM-DC is slightly increased due to the additional active period for downstream traffic. When the arrival traffic rate is high, the awake time increases and close to the level as in the UCS case. With respect with energy efficiency, the proposed energy management mechanism outperforms the traditional power ignoring mechanism. Particularly, using the uplink centric scheduling scheme saves up to 90% power on time. The energy consumption in the DCS case varies when the traffic load changes.

Figure 5.9 shows the network throughput under the PIS and EMM cases. The downstream throughput is monitored, because the upstream throughput is not affected by the proposed power saving mechanism. When the EMM is deployed, both DCS and UCS schemes degrade the throughput performance. In the DCS case, throughput is decreased due to the additional overhead by introducing the sleep mode GATE control message. Because the wakeup time and sleep time are assigned to ONUs when they receive every downstream packet, it causes a portion of control message overhead. In the UCS case, the network throughput is reduced greatly. The reason is that the OLT can dispatch downstream

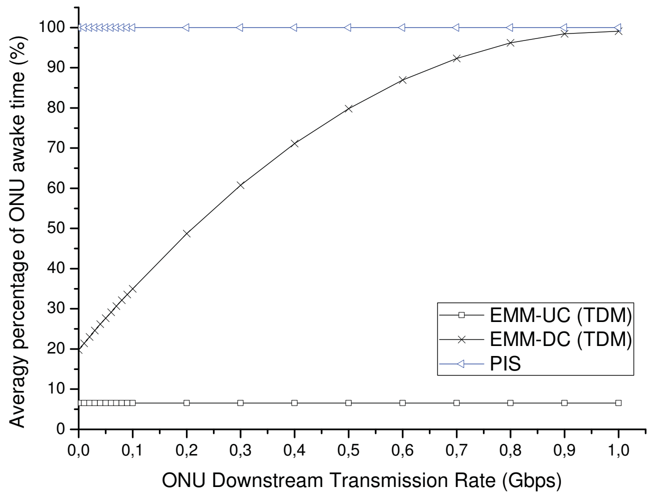

Figure 5.8: Average ONU awake time in PIS, EMM-UC, and EMM-DC algorithms.

traffic during its allocated upstream transmission period. Assuming the OLT polls and grants all ONUs in a Round Robin (RR) order, each ONU can derive one over sixteen percent of total downstream link capacity. The link utilization factor, computed as the amount of traffic delivered to the destination successfully over the total link capacity in the simulation time. The higher the network throughput, the high link utilization. The proposed EMM receives similar results compared to the PIS case, when the DCS scheme is applied. However, performance is highly downgraded in the UCS scheme.

The average queueing delay for analyzing performance differences under different power saving settings is shown in Figure 5.10. The average delay of the EMM DC scheme is better than that of the UCS scheme since, with UCS scheme, downstream packets have to wait for their awake period in the OLT buffer. We observe that when the traffic intensity is low, average delay remains constant since the transmission rate is high enough to serve under a range of values of traffic load. However, when the incoming traffic intensity reaches a certain point and causes congestion, average delay increases rapidly. As expected, the

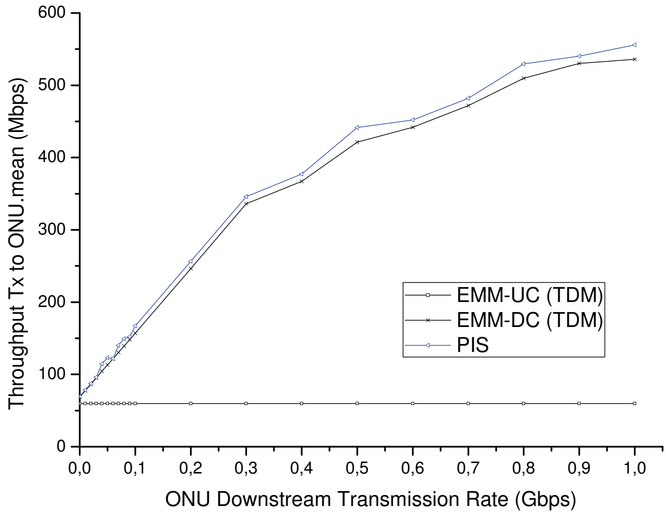

Figure 5.9: Network throughput in PIS, EMM-UC, and EMM-DC algorithms.

average queueing delay is consistent with the throughput performance, where the EMM DC scheme achieves better performance.

5.5.3 Impact of Upstream Bandwidth Allocation Schemes

Now we examine the impact of upstream bandwidth allocation schemes on the proposed energy management mechanism with the uplink centric scheme. In this scenario, only upstream traffic from ONUs to the OLT is considered to highlight the results differed from different upstream bandwidth allocation schemes. If the downstream traffic is involved, then additional time is required for transmitting downstream packets.

Figure 5.11 illustrates a simple scenario that shows the relationship between the upstream bandwidth allocation algorithm, TDM and DBA, with the ONU awake time. Because the downstream traffic is ignored, the sleep period is assigned to each ONU only based on the allocated upstream bandwidth, as the EMM UCS performs. In the TDM case, each ONU is assigned with a fixed slot time. In the DBA scheme case,

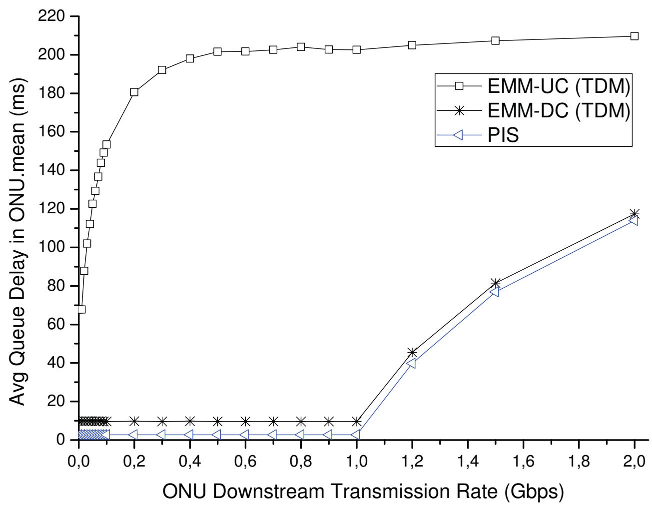

Figure 5.10: Average queueing delay in PIS, EMM-UC, and EMM-DC algorithms.

the OLT grants ONUs according to their bandwidth requests. Under low traffic load, buffered data at ONUs are small, and therefore, small amount of requested time slots. The percentage of awake time is increased along with the increased traffic rate. Because the maximum upstream slot time is constrained as 2 ms, the DBA based approach has similar results under the heavy traffic load.

Figure 5.12 and Figure 5.13 show the network throughput and queuing delay for the downstream traffic, respectively. The volume of downstream traffic is fixed at 1 Gbps and destination is uniformly distributed among 16 ONUs. Consistent with the curve of awake time, the downstream throughput increases when the ONU receives high input traffic and requests for large timeslots. On the other hand, when the required upstream transmission period is small, ONUs are turned off most of the observation time and downstream data are queued at the OLT. Thus the queuing delay is high under light traffic load in the dynamic bandwidth allocation case.

From previous studies [37], dynamic bandwidth allocation schemes improve EPON channel utilization and provide QoS guarantee. Our sim-

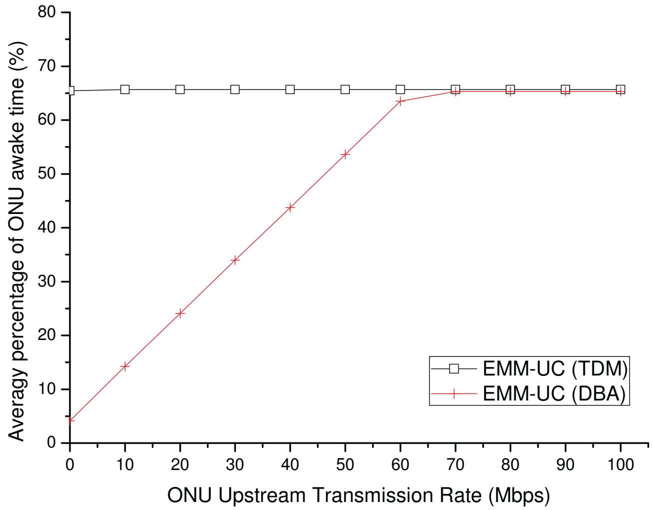

Figure 5.11: Awake time compared in different upstream bandwidth allocation schemes.

ulations over the upstream channel have demonstrated that the dynamic upstream bandwidth allocation scheme gains system power efficiency. However, the delay performance is downgraded when the upstream traffic load is low.

5.5.4 Impact of the sleep mode GATE message overhead

In this study, simulations show impacts of the sleep mode GATE control message. The MPCP control messages, GATE and REPORT, are formatted according to the IEEE 802.3ah specification. The sleep mode GATE message, G_s, is an extension of the standard GATE message. The additional two 4 bytes fields are embedded as assigned sleep time and wakeup time. The total size of G_s message is set to 41 bytes. In the UCS case, the sleep mode GATE message is used to poll ONUs, grant upstream bandwidth, and inform assigned sleep period. The Gs is generated and sent to ONUs in the same means as the original MPCP case. However, in the DCS case, G_s message is transmitted together with

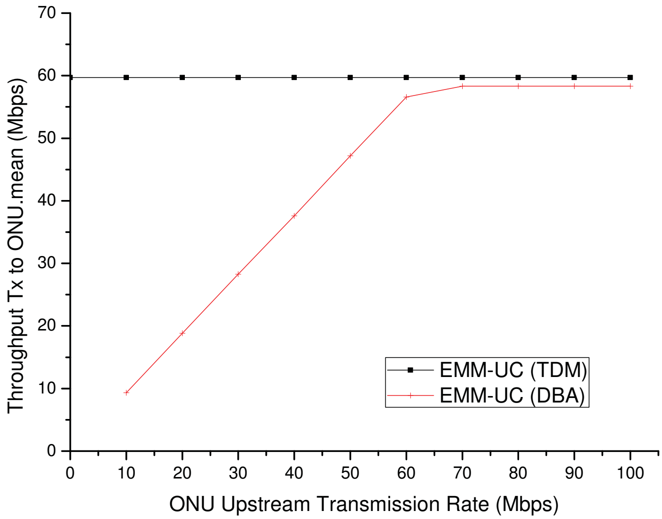

Figure 5.12: Network throughput compared in different upstream bandwidth allocation schemes.

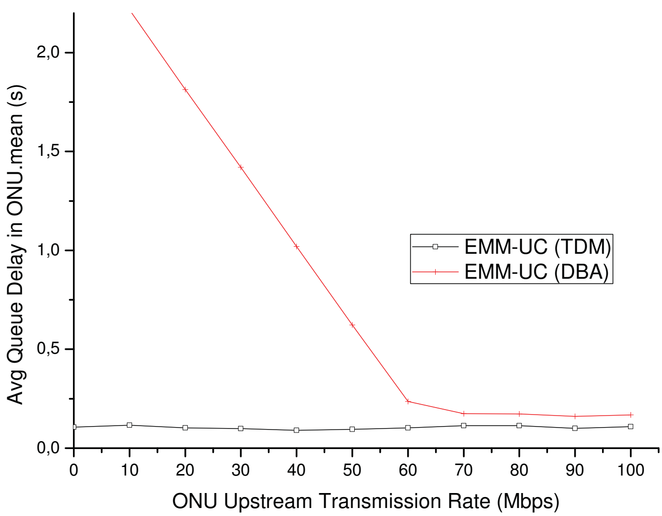

Figure 5.13: Average queuing delay compared in different upstream bandwidth allocation schemes.

each downstream packet, which results in amount of control message overhead. Therefore, in this test case, we simulate a network scenario with 16 ONUs and focus on EMM DCS scheme.

Figure 5.14 shows the number of transmitted G_s as a function of incoming packet sizes. In all cases, the total number of G_s is increased at the downstream data rate increases. The size of data packets are varied between 400 bits and 9400 bits. Under a certain input data rate, the larger the data size, the less transmitted G_s control message. With respect to bandwidth utilization, increased number of G_s message introduces high value of control overhead, therefore, results in low bandwidth utilization.

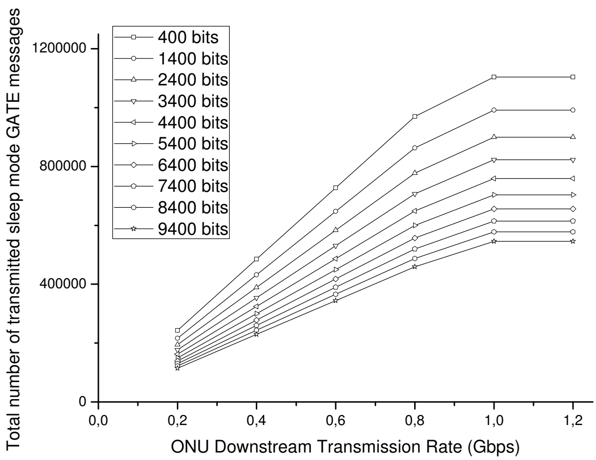

Figure 5.14: the number of transmitted G_s as a function of incoming packet sizes.

The delay performance is illustrated by Figure 5.15. The delay difference is hard to tell when the downstream traffic load is low. When traffic load becomes high, the large packet size case produces better delay performance. This is for the reason that higher channel utilization can be achieved by reducing the ratio of the overhead of G_s message and the size of payload.

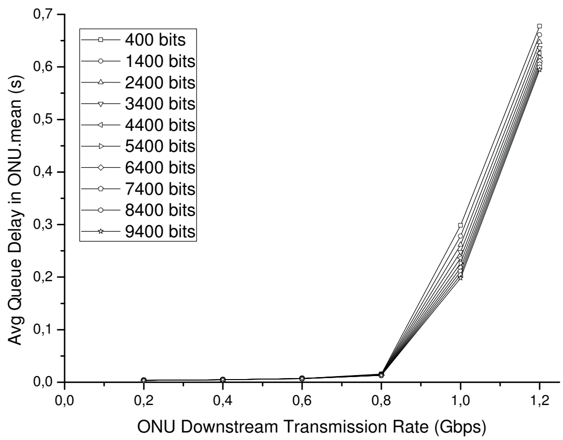

Figure 5.15: Average queuing delay performance as an impact of the sleep mode GATE message overhead.

5.5.5 Performance comparison with multiple ONUs

We now consider a scenario where 32 ONUs are connected to the OLT. With this network structure, an ONU is allocated less upstream transmission periods. Figure 5.16 shows the awake periods are reduced in both UCS and DCS schemes compared to the 16 ONUs structure. Consequently, the queuing delay is increased (shown in Figure 5.17). When the traffic load becomes high enough, the downstream traffic takes all transmission time and consumes the same power as in the PIS case.

5.6 Summary

In this chapter, we focus on the energy consumption in the EPON system. We proposed and investigated an analytical model for the energy expenditure in EPON. We developed an energy management mechanism for optimal assignments of sleep periods to ONUs. The delay constraints that need to be satisfied is analyzed in the scheduling problem. We solve the problem by proposing two heuristic scheduling algorithms:

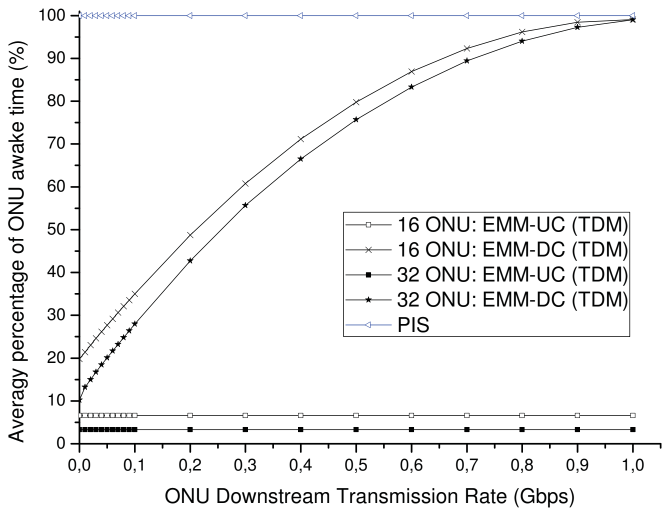

Figure 5.16: Awake time compared in scenarios with different connected AGs.

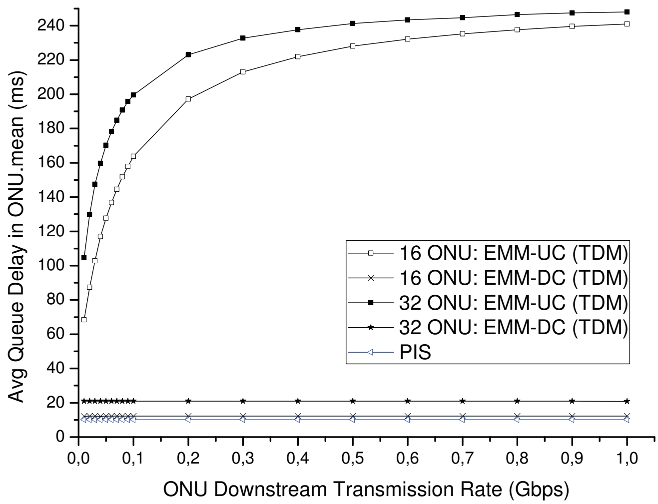

Figure 5.17: Network throughput compared in scenarios with different connected AGs.

the upstream centric scheduling algorithm and the downstream centric scheduling algorithm. An EPON system with energy efficiency based scheduling schemes are simulated in the OPNET simulator. We verified the solution quality with respect to a set of chosen metrics such as awake time and queuing delay. Specifically, we compared the energy performance and the network performance under different traffic load and different upstream bandwidth allocation schemes. We found that the energy management mechanism reduces the power consumption, and the DCS algorithm outperforms the UCS algorithm in the delay performance.

Chapter 6

Conclusions and Future Research

This dissertation has investigated the integrated control platform and made four important contributions to the design and modeling of integrated optical wireless networks. In this chapter, the content of this work is summarized at first, and then suggestions for future research are provided.

6.1 Summary of Thesis

The objective of this research is to develop and design an integrated control platform which provides bandwidth efficiency, power efficiency, and at the same time guarantees QoS requirements. Compared to the traditional simply connected hybrid network architecture, the integrated control platform is meant to manage network resources according to an overview and cooperation of both optical and wireless domains. these goals are achieved by proposing new functions, extending the legacy control system, obtaining the solution by theoretical analysis, and evaluating performances by simulations.

In Chapter 2, a converged optical wireless network architecture is introduced and an integrated EPON and WiMAX network model is presented. Motivations for introducing a joint and cooperative control platform are addressed. Functions of the integrated control platform are outlined and main research topics are addressed. The call admission control algorithm, packet scheduling schemes, load balancing mechanism, and energy efficiency scheme are discussed. The algorithms to optimize the integrated network performances by jointly designing control panels including both optical and wireless networks are reviewed.

Chapter 3 investigates the admission control scheme for the converged EPON and WiMAX network scenario. A new scheme, Integrated Optical Wireless Admission Control (IOW-AC), has been proposed to provide QoS guarantee. The IOW-AC is favorable in providing QoS satisfactions of real-time applications and improve the throughput of the best effort traffic. The performance differences between the IOW-AC and traditional AC schemes has been discussed and explained. Results have shown improvement in the utilization of network resources and queue sizes of the access gateway as well. To the best of our knowledge, this is the first traffic management solution that jointly considers QoS in both the optical domain and the wireless domain and optimizes the usage of network resources. The analysis can be a good assistance or guidance in future researches in many fields of hybrid optical wireless network architecture, including resource management, optimizing QoS and demanding service provisioning.

In Chapter 4, the load balancing problem is tackled in Integrated Optical Wireless Networks, where cell breathing technique is used to solve congestion by changing the coverage area of a fully loaded cell. It is aiming to be a load balancing mechanism to minimize network congestion and maximize overall network throughput. An efficient load balancing mechanism based on exchange of information between optical networks and wireless networks is proposed. Specifically, centralized power assignment and traffic allocation are jointly considered in the access gateways at the border of hybrid network architecture. Two alternative feedback schemes are proposed to report wireless network status. The performance of the proposed cell breathing in integrated EPON and WiMAX system is evaluated in terms of network throughput, packet dropping probability, and average queuing delay. Simulation results have shown that cell breathing significantly outperforms fixed power scheme.

In Chapter 5, the energy consumption issue has been studied for an EPON model. The main idea is to put ONUs into the sleep mode and determine a suitable wakeup time schedule at the OLT. The extensions to the traditional MPCP control messages are introduced and the energy consumption equations are formulated in both downstream scheduling cases. Two types of downstream scheduling schemes, Upstream Centric Scheduling (UCS) and Downstream Centric Scheduling (DCS), are proposed and compared. They are different in the way of assigning power 'active' and 'sleep' states to ONUs. There is a trade-off decision between maximizing the power saving and guaranteeing the network performance at the same time. Simulation results in terms of energy consumption and queuing delay are shown for the power saving enabled EPON system. It has proved that the energy management mechanism could reduce the power consumption at the cost of increased queuing delay.

6.2 Directions for Future Research

Convergence of wired and wireless networks is a trend of broadband access network architectures and benefits operators in terms of more connected users, higher bandwidth, and more applications. The topic of integrated control platform design is a timely research topic and our work provides good guidance and references for future research. Brief pointers for potential future research in this field are provided.

First, open issues such as mobility and resilience in the integrated control functions. Secondly, the existing protocols in both EPON and WiMAX systems have been investigated and the enhanced approaches are proposed. A design of a unified control platform from scratch, specifically for the integrated network architecture, could be considered. At last, the basic design ideas, design methodology, and analysis presented so far can be applied for investigating other converged optical and wireless networks, for instance, Gigabit Passive Optical Networks (GPON) and Wavelength Division Multiplexing Passive Optical Networks (WDM-PON) in the optical domain, and Wireless LAN (WLAN) and cellular systems in the wireless networks.

Bibliography

[1] Y. Yan, H. Yu, and L. Dittmann, "Wireless channel condition aware scheduling algorithm for hybrid optical/wireless networks," in *3rd International Conference on Access Networks*, 2008.

[2] Y. Yan, H. Yu, H. Wang, and L. Dittmann, "Integration of EPON and WiMAX networks: Uplink scheduler design," in *SPIE Symposium on Asia Pacific Optical Communications*, 2008.

[3] Y. Yan and L. Dittmann, "Enhanced signaling scheme admission control in the Hybrid Optical Wireless (HOW) networks," in *High-Speed Networks Workshop, co-located with IEEE Infocom conference*, 2009.

[4] Y. Yan, H. Yu, H. Wessing, and L. Dittmann, "Integrated resource management for Hybrid Optical Wireless (HOW) networks," in *International Conference on Communications and Networking in China (ChinaCom)*, 2009.

[5] S. Wong, Y. Yan, L. Kazovsky, and L. Dittmann, "MPCP assisted power control and performance of cell breathing in integrated EPON-WiMAX network," in *IEEE Global Communications Conference (GLOBECOM)*, 2009.

[6] Y. Yan, L. Dittmann, S. Wong, and L. Kazovsky, "Integrated control platform with load balancing algorithm in hybrid optical wireless networks," in *Third International Conference on New Technologies, Mobility and Security (NTMS)*, 2009.

[7] Y. Yan, H. Yu, H. Wessing, and L. Dittmann, "Integrated resource management framework in hybrid optical wireless networks," *IET*

Optoelectronics Special Issue on Next Generation Optical Access, 2009.

[8] Y. Yan, S. Wong, L. Valcarenghi, S. Yen, D. Campelo, S. Yamashita, L. Kazovsky, and L. Dittmann, "Energy management mechanism for Ethernet Passive Optical Networks (EPONs)," in *IEEE International Conference on Communications (ICC)*, 2010.

[9] Y. Yan, L. Dittmann, S. Wong, and L. Kazovsky, "Load balancing in integrated optical wireless networks: Algorithms and evaluation," in *European Wireless (EW2010)*, 2010.

[10] Y. Yan and L. Dittmann, "Modeling energy management mechanism in ethernet passive optical networks," in *13th Communications and Networking Simulation Symposium (CNS'10)*, 2010.

[11] L. Kazovsky, S. Wong, V. Gudla, P. Afshar, S. Yen, S. Yamashita, and Y. Yan, "Addressing reach extension and reliability weaknesses in next generation optical access networks: A survey," *IET Optoelectronics Special Issue on Next Generation Optical Access*, 2010.

[12] H. Yu, Y. Yan, and M. Berger, "IPTV traffic management in Carrier Ethernet transport networks," in *OPNETWORK 2008*, 2008.

[13] H. Yu, Y. Yan, and M. Berger, "IPTV traffic management using topology-based hierarchical scheduling in Carrier Ethernet transport networks," in *International Conference on Communications and Networking in China (ChinaCom)*, 2009.

[14] H. Yu, Y. Yan, and M. Berger, "Topology-based hierarchical scheduling using deficit round robin: Flow protection and isolation for triple play service," in *International Conference on Future Information Networks*, 2009.

[15] H. Wessing, Y. Yan, and M. Berger, "Modelling direct application to network bandwidth provisioning for high demanding research applications," in *5th International Conference on Applied Informatics and Communications (AIC'05)*, 2005.

[16] H. Wessing, Y. Yan, and M. Berger, "Simulation bases analysis on dynamic resource provisioning in optical networks using GMPLS

technologies," *WSEAS Transaction on Communications*, vol. 5, pp. 9–16, 2006.

[17] H. Wessing, Y. Yan, and M. Berger, *Definition of application needs and scale of dynamics in research network infrastructures - additional deliverable on modelling topics*. MUPBED - Multi-Partner European Test Beds for Research Networking Deliverable D2.1a, 2006.

[18] Y. Yan, H. Wessing, M. Berger, and L. Dittmann, "Prioritized OSPF-TE mechanism for multimedia applications in MPLS networks," in *5th G/MPLS Networks conference*, 2006.

[19] Y. Yan, H. Wessing, and M. Berger, "Bidirectional RSVP-TE for multimedia applications in GMPLS networks," in *10th International Conference on Optical Network Design and Modelling*, 2006.

[20] L. Kazovsky, W. Shaw, D. Gutierrez, N. Cheng, and S. Wong, "Next-generation optical access networks," *Journal of Lightwave Technology*, vol. 25, pp. 3428–3442, 2007.

[21] Ethernet in the First Mile Alliance (EFMA), *Ethernet Passive Optical Network (EPON) tutorial, Revision 4*, 2004.

[22] ITU-T Recommendation G.984.1, *Gigabit Capable Passive Optical Networks (GPON): General Characteristics*, 2003.

[23] J. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX*. Prentice Hall, 2007.

[24] IEEE 802.11 Working Group for WLAN Standards, *http://ieee802.org/11/*.

[25] The IEEE 802.16 Working Group on Broadband Wireless Access Standards, *http://ieee802.org/16/*.

[26] C. Youjun, H. Tian, L. Sun, and X. Haibo, "Research on the acces network and MAC technique for beyond 3G systems," *IEEE Wireless Communications*, vol. 14, pp. 57–61, 2007.

[27] W. Shaw, D. Gutierrez, K. S. Kim, N. Cheng, S. Wong, S. Yen, and L. Kazovsky, "GROW-Net a hybrid optical wireless access network," in *9th Joint Conf. Information Sciences*, 2006.

[28] S. Sarkar, S. Dixit, and B. Mukherjee, "Hybrid wireless-optical broadband-access network (WOBAN): a review of relevant challenges," *Journal of Lightwave Technology*, vol. 25, pp. 3329–3340, 2007.

[29] G. Shen, R. Tucker, and C. Chae, "Fixed mobile convergence architectures for broadband access: integration of EPON and WiMAX," *IEEE Communications Magazine*, vol. 45, pp. 44–50, 2007.

[30] S. Sarkar, *Design and Analysis of Wireless-Optical Broadband Access Networks (WOBAN).* PhD thesis, University of California Davis, 2008.

[31] L. Kazovsky, N. Cheng, W. Shaw, and S. Wong, "CROWN - converged optical and wireless networks: network architecture and routing algorithms," in *International Conference on Transparent Optical Networks*, pp. 266–269, 2007.

[32] EURESCOM P816 - Implementation frameworks for integrated wireless - optical access networks, *http://www.eurescom.eu/ pub-deliverables/P800-series/P816/*, 2000.

[33] M. McGarry, M. Maier, and M. Reisslein, "Ethernet PONs: a survey of dynamic bandwidth allocation (DBA) algorithms," *IEEE Communications Magazine*, vol. 42, pp. 8–15, 2004.

[34] Y. Luo, S. Yin, N. Anson, and T. Wang, "Resource management for broadband access over time-division multiplexed passive optical networks," *IEEE Network*, vol. 21, pp. 20–27, 2007.

[35] WiMAX Forum, *http://www.wimaxforum.org/*.

[36] IEEE 802.3ah Ethernet in the First Mile Task Force, *http://www.ieee802.org/3/efm/*.

[37] K. Glen, M. Biswanath, and P. Gerry, "IPACT: a dynamic protocol for an Ethernet PON (EPON)," *IEEE Communication Magazine*, vol. 40, pp. 74–80, 2002.

[38] C. Assi, Y. Yinghua, S. Dixit, and M. Ali, "Dynamic bandwidth allocation for quality-of-service over Ethernet PON," *IEEE Journal on Selected Areas in Communications*, vol. 21, pp. 1467–1477, 2003.

[39] M. Ma, Y. Zhu, and T. H. Cheng, "A systematic scheme for multiple access in Ethernet passive optical access networks," *Journal of Lightwave Technology*, vol. 23, pp. 3671–82, 2005.

[40] C. Cicconetti, A. Erta, L. Lenzini, and E. Mingozzi, "Performance evaluation of the IEEE 802.16 MAC for QoS support," *IEEE Transactions on Mobile Computing*, vol. 6, pp. 26–38, 2007.

[41] D. Niyato and E. Hossain, "Queue-aware uplink bandwidth allocation and rate control for polling service in IEEE 802.16 broadband wireless networks," *IEEE Transactions on Mobile Computing*, vol. 5, pp. 668–679, 2006.

[42] J. Yu, "Scheduling and performance analysis of QoS for IEEE 802.16 broadband wireless access networks," *A chapter in WiMAX Handbook (published by Taylor & Francis Group)*, pp. 57–75, 2007.

[43] A. Sayenko, O. Alanen, and T. Hamalainen, "Scheduling solution for the IEEE 802.16 base station," *Journal of Computer Networks*, vol. 52, pp. 96–115,, 2008.

[44] A. Lera, A. Molinaro, and S. Pizzi, "Channel-aware scheduling for QoS and fairness provisioning in IEEE 802.16/WiMAX broadband wireless access systems," *Journal of IEEE Network*, vol. 21, pp. 34–41, 2007.

[45] X. Bai, A. Shami, and Y. Ye, "Robust qos control for single carrier PMP mode IEEE 802.16 systems," *Journal of IEEE Transactions on Mobile Computing*, vol. 7, pp. 416–429, 2008.

[46] B. Rong, Y. Qian, and H. chen, "Adaptive power allocation and call admission control in multiservice WiMAX access networks," *Journal of IEEE Wireless Communications*, vol. 14, pp. 14–19, 2007.

[47] B. Chang and Y. Chen, "Adaptive hierarchical polling and markov decision process based CAC for increasing network reward and re-

ducing average delay in IEEE 802.16 WiMAX networks," *Journal of Computer Communications*, vol. 31, pp. 2280–2292, 2008.

[48] D. Niyato and E. Hossain, "Qos-aware bandwidth allocation and admission control in IEEE 802.16 broadband wireless access networks: A non-cooperative game theoretic approach," *Journal of Computer Networks*, vol. 51, pp. 3305–3321, 2007.

[49] M. Sauer, A. Kobyakov, and J. George, "Radio over fiber for picocellular network architectures," *Journal of Lightwave Technology*, vol. 25, pp. 3301–3320, 2007.

[50] W. T. Shaw, S. W. Wong, N. Cheng, K. Balasubramanian, X. Zhu, M. Maier, and L. Kazovsky, "Hybrid architecture and integrated routing in a scalable optical-wireless network," *IEEE/OSA Journal of Lightwave Technology*, vol. 25, pp. 3343–3351, 2007.

[51] S. Sarkar, H. Yen, S. Dixit, and B. Mukherjee, "Dara: delay-aware routing algorithm in a hybrid wireless-optical broadband access network (WOBAN)," in *IEEE International Conference on Communications*, pp. 2480–2484, 2007.

[52] H. Zhang, "Service disciplines for guaranteed performance service in packet-switching networks," *Proceedings of the IEEE*, vol. 83, pp. 1374–1396, 1995.

[53] H. Fattah and C. Leung, "An overview of scheduling algorithms in wireless multimedia networks," *IEEE Wireless Communications*, vol. 9, pp. 76–83, 2002.

[54] H. Kim and A. Wolisz, "A radio over fiber based wireless access network architecture for rural areas," in *14th IST Mobile and Wireless Communications Summit*, 2005.

[55] P. Lin, C. Qiao, T. Wang, and J. Hu, "Optimal utility-based bandwidth allocation over integrated optical and WiMAX networks," in *Optical Fiber Communication Conference and National Fiber Optic Engineers Conference*, p. 3, 2006.

[56] K. Yang, S. Ou, K. Guild, and H. Chen, "Convergence of ethernet PON and IEEE 802.16 broadband access networks and its QoS-aware dynamic bandwidth allocation scheme," *IEEE Journal on Selected Areas in Communications*, vol. 27, pp. 101–116, 2009.

[57] W. Song, W. Zhuang, and Y. Cheng, "Load balancing for cellular/WLAN integrated networks," *IEEE Network*, vol. 21, pp. 27–33, 2007.

[58] B. Sikdar, "An analytic model for the delay in IEEE 802.11 PCF MAC-based wireless networks," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 1542–1550, 2007.

[59] Y. Luo, T. W. S. Yin, Y. Suemura, S. Nakamura, and M. Cvijectic, "QoS-aware scheduling over hybrid optical wireless networks," in *Optical Fiber Communication Conference and National Fiber Optic Engineers Conference*, pp. 1–7, 2006.

[60] OPNET Modeler 14.5, *http://www.opnet.com/*, 2005.

[61] A. Balachandran, G. M. Voelker, P. Bahl, and V. Rangan, "Characterizing user behavior and network performance in a public wireless LAN," *ACM SIGMETRICS*, vol. 30, pp. 195–205, 2002.

[62] M. Balazinska and P. Castro, "Characterizing mobility and network usage in a corporate wireless local-area network," in *MobiSys '03: Proceedings of the 1st international conference on Mobile systems, applications and services*, pp. 303–316, 2003.

[63] T. Alpcan, X. Fan, T. Basar, M. Arcak, and J. T. Wen, "Power control for multicell CDMA wireless networks: A team optimization approach," *Wireless Networks*, vol. 14, pp. 647–657, 2008.

[64] A. Ghasemi and E. Sousa, "Jointly optimal base station assignment for up- and downlink CDMA through space division duplex," in *IEEE Wireless Communications and Networking Conference*, pp. 2226–2231, 2004.

[65] S. Kiani and D. Gesbert, "Optimal and distributed scheduling for multicell capacity maximization," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 288–297, 2008.

[66] Y. Bian and A. Nix, "Mobile WiMAX: Multi-cell network evaluation and capacity optimization," in *IEEE Vehicular Technology Conference*, pp. 1276–1280, 2008.

[67] S. Hanly, "An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1332–1340, 1995.

[68] C. Oh and A. Yener, "Downlink throughput maximization for interference limited multiuser systems: TDMA versus CDMA," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 2454–2463, 2007.

[69] S. Kandukuri and S. Boyd, "Optimal power control in interference-limited fading wireless channels with outage-probability specifications," *IEEE Transactions on Wireless Communications*, vol. 1, pp. 46–55, 2002.

[70] H. Kim and A. Wolisz, "Load balancing for centralized wireless networks," in *16th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1875–1879, 2005.

[71] N. Bambos and S. Kandukuri, "Power controlled multiple access (PCMA) in wireless communication networks," in *IEEE INFOCOM*, pp. 386–395, 2000.

[72] J. Qiu and J. Mark, "A dynamic load sharing algorithm through power control in cellular CDMA," in *the Ninth IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 3, pp. 1280–1284, 1998.

[73] L. Du, J. Bigham, and L. Cuthbert, "A bubble oscillation algorithm for distributed geographic load balancing in mobile networks," in *IEEE INFOCOM*, vol. 1, p. 9, 2004.

[74] P. Bahl, M. T. Hajiaghayi, K. Jain, S. V. Mirrokni, L. Qiu, and A. Saberi, "Cell breathing in wireless LANs: Algorithms and evaluation," *IEEE Transactions on Mobile Computing*, vol. 6, pp. 164–178, 2007.

[75] A. Sang, X. Wang, M. Madihian, and R. Gitlin, "Coordinated load balancing, handoff/cell-site selection, and scheduling in multi-cell packet data systems," in *the Tenth Annual International Conference on Mobile Computing and Networking*, pp. 302–314, 2004.

[76] IEEE 802.16e, *IEEE standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems*, 2005.

[77] W. Joseph and L. Martens, "Performance evaluation of broadband fixed wireless system based on IEEE 802.16," in *IEEE Wireless Communications and Networking Conference*, pp. 978–983, 2006.

[78] T. Ikeda, M. Asa, and K. Saito, "Comments on IEEE 802.16j path-loss models in IEEE802.16j-06/013," in *IEEE 802.16's Relay Task Group*, pp. 978–983, 2006.

[79] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[80] S. Murawwat and K. Ahmed, "Performance analysis of 3G and WiMAX as cellular mobile technologies," in *Second International Conference on Electrical Engineering (ICEE)*, pp. 1–5, 2008.

[81] P. Mach and R. Bestak, "WiMAX performance evaluation," in *IEEE 6th International Conference on Networking*, pp. 91–94, 2007.

[82] C. Lange and A. Gladisch, "Energy consumption of telecommunication networks: A network operator's view," in *Work shop on Optical Fiber Communication Confence (OFC) "Energy Footprint of ICT: Forecasts and network solutions"*, 2009.

[83] C. Lange, D. Kosiankowski, C. Gerlach, F. Westphal, and A. Gladisch, "Energy consumption of telecommunication networks," in *35th European Conference on Optical Communication (ECOC)*, pp. 1–2, 2009.

[84] M. Gupta and S. Singh, "Greening of the internet," *ACM SIGCOMM*, pp. 19–26, 2003.

[85] J. Baliga, R. Ayre, K. Hinton, W. V. Sorin, and R. S. Tucker, "Energy consumption in optical IP networks," *IEEE/OSA Journal of Lightwave Technology (JLT)*, vol. 27, pp. 2391–2403, 2009.

[86] K. Christensen, C. Gunaratne, B. Nordman, and A. George, "The next frontier for communications networks: Power management," *Computer Communications*, vol. 27, pp. 1758–1770, 2004.

[87] C. E. Jones, K. M. Sivalingam, P. Agrawal, and J. C. Chen, "A survey of energy efficient network protocols for wireless networks," *Wireless Networks*, vol. 7, pp. 343–358, 2001.

[88] J. Lee, C. Rosenberg, and E. Chong, "Energy efficient schedulers in wireless networks: Design and optimization," *Mobile Networks and Applications*, vol. 11, pp. 377–389, 2006.

[89] Y. Zhang, "Performance modeling of energy management mechanism in IEEE 802.16e mobile WiMAX," in *Wireless Communications and Networking Conference (WCNC)*, pp. 3205–320, 2007.

[90] R. Cohen and B. Kapchits, "An optimal wake-up scheduling algorithm for minimizing energy consumption while limiting maximum delay in a mesh sensor network," *IEEE/ACM Transactions on Networking*, vol. 17, pp. 570–581, 2009.

[91] J. Hsieh, T. Lee, and Y. Kuo, "Energy-efficient multi-polling scheme for wireless LANs," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 1532–1541, 2009.

[92] Y. Chen and S. Tsao, "Energy-efficient sleep-mode operations for broadband wireless access systems," in *IEEE Vehicular Technology Conference*, pp. 1112–1116, 2006.

[93] T. Smith, R. S. Tucker, K. Hinton, and A. V. Tran, "Implications of sleep mode on activation and ranging protocols in PONs," *21st Annual Meeting of the IEEE Lasers and Electro-Optics Society (LEOS)*, vol. 11, pp. 604–605, 2008.

[94] S. Wong, L. Valcarenghi, S. Yen, D. Campelo, S. Yamashita, and L. Kazovsky, "Sleep mode for energy saving PONs: Advantages

and drawbacks," in *Globlecom'09 Second International workshop on Green Communication*, 2009.

[95] R. Kubo, J. Kani, Y. Fujimoto, N. Yoshimoto, and K. Kumozaki, "Proposal and performance analysis of a power-saving mechanism for 10 gigabit class passive optical network systems," in *NOC*, vol. 1, pp. 87–94, 2009.

[96] IEEE 802.3az Meeting, *EPON Powersaving via Sleep Mode*, 2008.

[97] I. Cerutti, L. Valcarenghi, and P. Castoldi, "Power saving architectures for unidirectional WDM rings," in *Optical Fiber Communication (OFC)*, pp. 1–3, 2009.